Ref. Ares(2019)4137066 - 30/06/2019

DataBench

Evidence-Based Big Data Benchmarking to Improve Business Performance

D2.3 Analysis of Actual and Emerging Industrial Needs

Abstract

One of the DataBench project's central goals is to design, develop and validate a benchmarking process based on highly relevant business metrics to help European organisations evaluate their use of Big Data and analytics technologies (BDTs) as they seek to improve their business performance.

This report presents further results of the economic and market analysis carried out in the context of Work Package 2 (WP2). The work documented in this report is based on the methodology described in the preceding deliverable (D2.1: Economic, Market and Business Analysis Methodology). This document provides further analysis of the performance measurement metrics that companies are benchmarking to assess their use of Big Data and analytics to that presented in the previous DataBench deliverable (D2.2: Preliminary Benchmarks of Industrial Significance of Big Data Technology Performance Parameters). The document is based on desk research from public sources, IDC research databases and a survey of 700 European businesses in 11 EU Member States. The report shows the importance of business KPIs for BDA users to benchmark the value of their BDA investments, analysed by sector and company size. These benchmarks will be tested and validated through the analysis of case studies carried out by WP4, verified through the analysis in WP5 and updated at the end of the project in 2020.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 780966.

Deliverable D2.3	Analysis of Actual and Emerging Industrial Needs
Work package	WP2
Task	2.3
Due date	30/06/2019
Submission date	June 30, 2019
Deliverable lead	IDC
Version	1.0
Authors	IDC (Gabriella Cattaneo, David Wells, Erica Spinoni) Polimi (Chiara Francalanci, Angela Geronazzo, Lucia Polidori, Barbara Pernici) ATOS (Tomas Pariente Lobo)
Reviewers	Polimi (Sergio Gusmeroli), ATOS (Nuria De Lama)

Keywords

Benchmarking, key performance indicators (KPIs), Big Data and analytics (BDA), Big Data technologies (BDTs), business performance, business report, economic impact, economic indicator, industrial relevance, performance metrics, self-assessment tool, technical performance tool, use cases.

Disclaimer

This document reflects the authors' views only. The European Commission is not responsible for any use that may be made of the information in this document.

Copyright Notice

Copyright belongs to the authors. Use of any materials from this document should be referenced and is at the user's own risk.

Table of Contents

Table of Figures

Table of Tables

Executive Summary

The overall objective of the DataBench project is to provide benchmarks and performance evaluation mechanisms to identify the business impact and industrial significance of the deployment of Big Data technologies in Europe.

Successful deployment of BDTs in the medium to long term depends on adopters' ability to convincingly benchmark and assess their impact on their businesses. This will primarily be based on the adopter's own business objectives, but also to enable comparison with and learning from best practice in their sector. This report is based on a survey of a representative sample of European business users, and provides in-depth analysis of users' needs in terms of benchmarking and assessment of BDT objectives and impacts.

Accordingly, the DataBench survey asked respondents which business KPIs they expected to improve through their existing or planned BDA investments. KPIs for which a significant impact is expected would be a suitable basis for benchmarking BDA impact. The survey also asked about respondents' current and planned BDT infrastructure.

This report examines in more detail the differences in how companies plan to benchmark their BDA investment based on their size and industry sector, to inform the development of benchmarking metrics by DataBench that are relevant for the European economy and industry. We also carried out qualitative and quantitative analysis of potential correlations between survey respondents' business expectations from BDA adoption and their current and planned analytics technical infrastructure. This is to inform a primary goal of DataBench — to identify how business benchmarks of actual or projected BDA benefits and technical benchmarks of actual or planned BDT capabilities can be aligned to maximise the value of BDA.

The main findings are:

1. Although there is a consensus in that all surveyed industry sectors expect BDA to contribute to a range of business KPIs, there are some significant differences in sector expectations that should inform appropriate BDA sector value benchmarks.
2. Averaged across all respondents, the cost reduction KPI was considered the least important measure of the value of BDT investments. This is surprising considering the wide range of BDA use cases that can make significant cost efficiency contributions, for example in supply chain optimisation across many sectors. It suggests that an understandable emphasis on the clear value of BDA as an enabler of more radical digital transformation may be overshadowing its ability to also make less spectacular but potentially equally valuable contributions to core business KPIs.
3. In contrast, SME respondents valued cost reduction as the most important KPI for measuring the value of their BDA adoption, and were significantly less interested in the positive impact that BDA might have on several other KPIs. Combined with the currently much lower level of BDA adoption among SMEs than among larger companies, this suggests that facilitating delivery of packaged BDA solutions suitable for SME adoption, and addressing a wider range of KPIs than only cost reduction, might lead to significant overall economic benefit.
4. A statistical cluster analysis of survey respondents' answers indicates that BDA adopters can be classified into five overlapping types, with a range of differing attitudes towards BDA adoption, from extremely pragmatic and business led to largely or entirely technology led. The design of effective business and technical

benchmarks should take account of these differing drivers and make it straightforward both for business-led adopters to deploy appropriate technical BDT benchmarks and for technology-led adopters to implement appropriate business benchmarks to measure, justify and optimise their BDT investment.

The rationales for these conclusions are detailed below.

User Needs by Industry

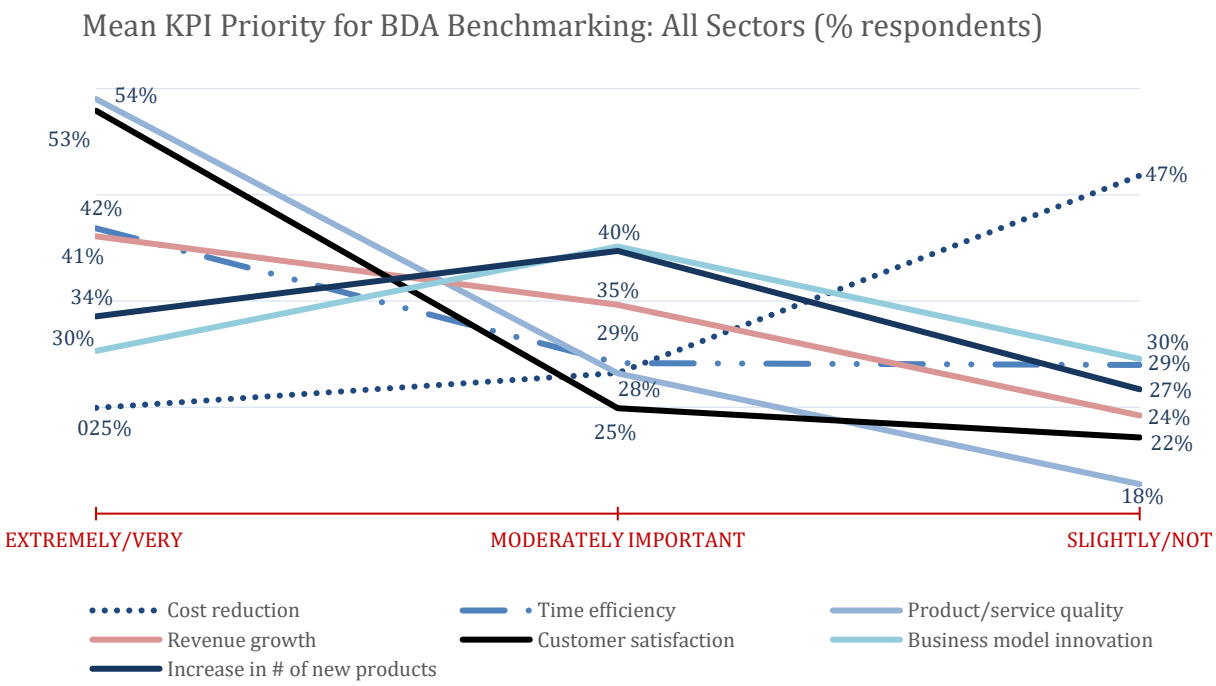


Figure 1 — Mean KPI Priority for BDA Benchmarking: All Sectors

Source: DataBench Survey, October 2018, 700 interviews

To establish best practice in benchmarking business impact on a per-sector basis, for this report we identified the sectors that had significantly higher or lower expectations that BDA would have an impact on specific KPIs than the averages across all respondents shown in the chart above.

We identified not only significant differences in KPI expectations between many of the sectors, but were able to informally group sectors, ranging from the conservative adopters (agriculture and healthcare) through to the business/IT services and telecom/media sectors, which expect BDA to contribute not only to many basic KPIs (e.g., customer satisfaction) but to more transformational KPIs such as business model innovation.

These sector specific KPI priorities should be considered by individual companies when assessing their own benchmarking strategy for BDA impact, and also taken into account by BDT vendors seeking to most effectively address sector priorities.

Although financial services, manufacturing and transport/logistics considered cost reduction more significant as a BDA KPI than other sectors, a somewhat surprising finding is that overall it is by far the least-favoured KPI, in all sectors, to assess BDA impact, compared with the other six KPIs in the DataBench model.

This may suggest that BDA adoption, especially if significant investment is required, is more easily justified internally to a company by appealing to its ability to boost less traditional KPIs. It also suggests perhaps that opportunities to deploy modern BDA to effectively address less innovative but equally important business goals such as cost efficiency may not always be being fully considered by adopters.

Business Size Factors and SME Needs

As reported in the previous DataBench deliverable (D2.2), the level of current and planned BDA adoption varies significantly by company size. The main difference is between SMEs, which significantly lag medium-sized and large enterprises in terms of adoption of BDA (with only 16% of them already using BDA), and large companies with over 1,000 employees (47% of which already use BDA). Given the vital role that SMEs play in the European economy, the low adoption of BDA must be a cause for concern.

The survey shows that for SMEs cost reduction is a significantly more important KPI for benchmarking BDA than for other company sizes. Product/service quality, revenue growth, customer satisfaction and business model innovation are less important for SMEs than for larger companies when assessing the impact of BDA.

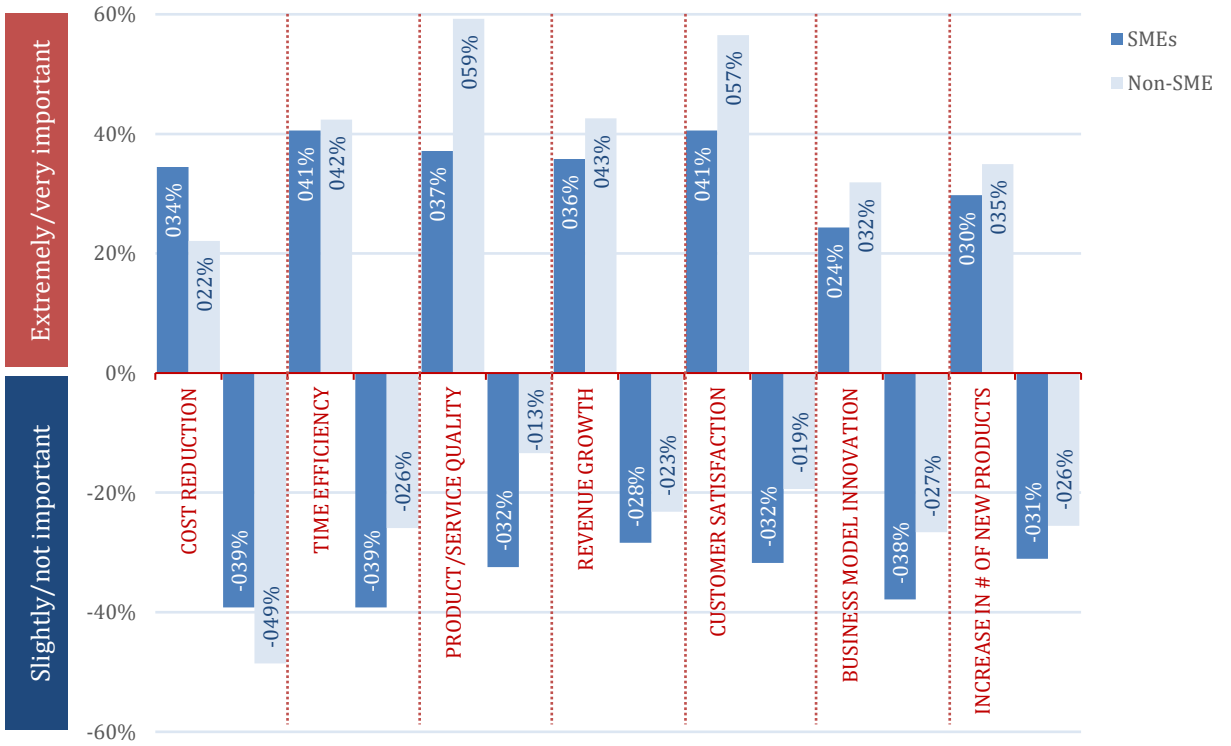


Figure 2 — Comparison of Stated Importance of KPIs in Benchmarking BDA Adoption Between SMEs and Non-SMEs

Source: DataBench Survey, October 2018, 700 interviews; positive bars are responses for "extremely/very important"; negative bars are for "slightly/not important"

While it is understandable that small companies must address cash flow/burn rate as a priority, there are numerous other valuable applications of BDA that could be of significant benefit to KPIs that the sector does not appear to consider as priorities for BDA adoption. This includes improving their understanding of the needs of their target market and making informed decisions about how best to meet them. There seems to be a significant opportunity for suppliers of BDT and BDA solutions to package these in ways that address the adoption obstacles of SMEs — to supply affordable, out-of-the-box analytics-based solutions that can be used without specialised skills and that draw on external data sources, such as environmental data or market research data, in ways that are relevant to the general or sector-specific business needs of SMEs.

Although some SMEs are already addressing more sophisticated BDA applications in a variety of sectors, there seems to be an opportunity for initiatives to encourage the BDA vendor community to more effectively address the market opportunities of SME BDA adoption and to encourage a similar market push from SMEs for usable and relevant analytics solutions. Facilitating BDA adoption by SMEs this way could bring significant additional economic benefit.

BDT Technology Choices by Type of Adopter

Qualitative and quantitative analysis was carried out to identify potential correlations between survey respondents' business expectations from BDA adoption and their current and planned future analytics technical infrastructure.

In the initial findings, we were surprised to see a slightly negative correlation between the current status of BD usage by adopters and the expected business benefits, and the level of investments in BDTs and the impact of BDA adoption on actual business benefits.

These, however, were subsequently explained by the identification of types of BDA adopter that are currently largely or entirely technology led and not presently concerned with measuring or even necessarily attempting to implement business benefits from BDTs. Such adopters are currently preparing highly capable BDT infrastructures to facilitate the rapid implementation of BDA use cases.

Five (to some extent overlapping) categories of BDA adopter were identified based on the statistical identification of common patterns in their survey responses. These have been informally labelled in approximate order from business led to technology led:

- Conservatives
- Optimists
- Promoters
- Explorers

Technology EnthusiastsThe development of DataBench business and technical benchmarks for BDA should take into account the differing attitudes to BDA adoption of companies aligned with these groups to ensure that business and technical benchmark adoption is straightforward, whatever the starting point and whoever the key stakeholders for the company's BDA adoption strategy happen to be.

1. Introduction

1.1. Objectives

One of the DataBench project's central goals is to design a benchmarking process based on highly relevant business metrics to help European organisations evaluate their use of Big Data technologies (BDTs) as they seek to improve their business performance. DataBench bridges the gap between technical and business benchmarking of BDTs, by investigating existing Big Data benchmarking tools and projects, identifying the main gaps, providing a robust set of metrics to compare technical results coming from those tools, and providing a framework to associate those technical results with key business use cases and economic processes.

WP2's goal is to carry out the economic and market analysis research, implementing two of the main objectives of the project:

- Objective II: "Performance of economic and market analysis to assess the 'European economic significance' of benchmarking tools and performance parameters"
- Objective III: "Evaluate the business impacts of BDT benchmarks of performance parameters of industrial significance"

This report builds on the results of the economic and market analysis presented in the D.2.2 deliverable (Preliminary Benchmarks of European and Industrial Significance) on the basis of statistical data from public sources and the DataBench survey of 700 European enterprises. The main goal of the report is to provide deeper analysis of Big Data technology industrial users and to contribute to the finalisation of benchmarks to be presented in the next deliverable, D.2.4: Benchmarks of European and Industrial Significance. More specifically the report aims to:

- Deepen the analysis of business KPIs of BDT by industry and use case
- Analyse the correlation between users' technical and business choices related to BDA adoption
- Present the plan for an extension of the DataBench survey with the industrial partners of the H2020 ICT 14 and 15 projects, which will be accompanied by a self-assessment tool, a real-time report providing respondents with a comparison between their KPIs and those of a group of their peers (same industry and company size) extracted from the previous survey results; this will help to enrich the KPI database and support the validation of the benchmarks

As shown in Figure 3, the KPIs and preliminary benchmarks presented in this report feed into the development of the benchmarking tool developed in WP3 and tested and validated in the in-depth case studies planned by WP4.

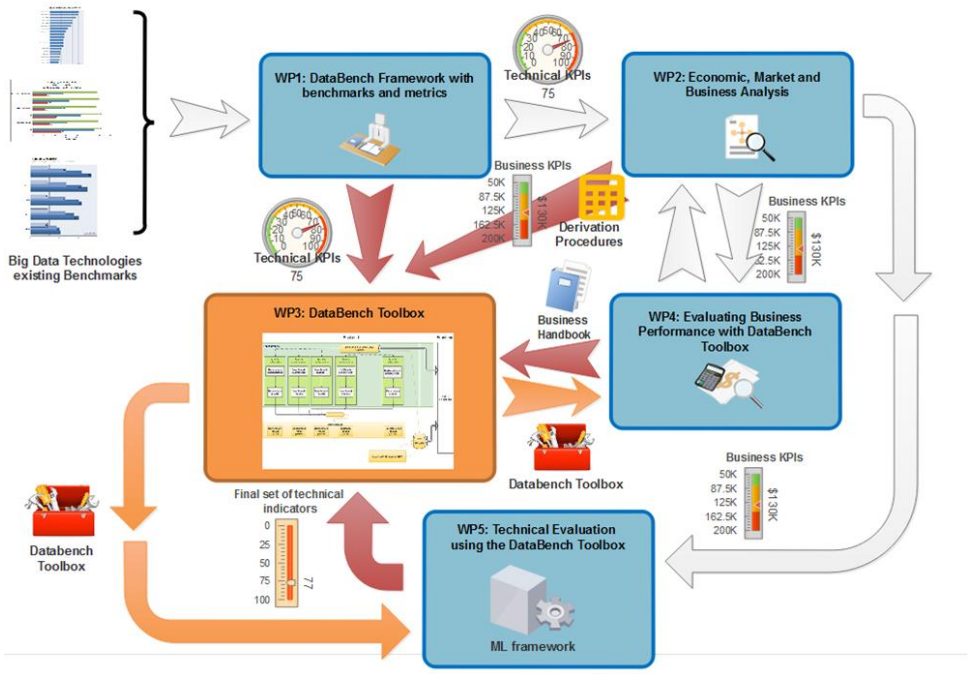


Figure 3 — DataBench Workflow and WP Roles

Building on this process, the project will finalise the industrial and business benchmarks (Deliverable 2.4 at month 24 of the project, December 2019) and will illustrate the utilisation of this benchmarking methodology in the final DataBench Benchmarking Handbook (D4.4 at M34). This research process is focused on developing useful tools for the BDT user community, for the Big Data Value (BDV) contractual Public-Private Partnership (cPPP) and the European Big Data industry.

The conceptual framework linking technical and business benchmarking and the approach to the design and selection of KPIs is presented in the first deliverable of WP1 (D.1.1: Industry Requirements, Benchmarking Metrics and KPIs) and has been made available in parallel with this report. This will also help to increase recognition and acceptance of the benchmarks by the industrial community.

1.2. Structure of the Report

The report is organised as follows:

- Chapter 1 presents the main objectives and methodological approach.
- Chapter 2 presents the analysis of users' needs by industry based on the relative relevance of KPIs.
- Chapter 3 presents the analysis of users' needs by company size, particularly concerning SMEs, based on the relative relevance of KPIs.
- Chapter 4 presents the H2020 ICT projects that will be targeted by DataBench in a second-wave survey.
- Chapter 5 presents the results of the descriptive and statistical analyses of technical choices.
- Chapter 6 draws the main conclusions of the report.

- The Annex includes:
  - A glossary of the main terms used in the report
  - A detailed description of the DataBench survey sample
  - Mapping of the H2020 ICT project trials and pilots
  - Examples of the new DataBench survey tool

1.3. Interdependencies with Other WPs

WP2 leads on the study of the business KPIs and economic impacts of BDTs through the survey analysis. The analysis and insights from WP2 are relevant inputs for other Work Packages, mainly WP3 and WP4. In fact, WP2 contributes to WP3, providing a basis for the business benchmarking to feed into the DataBench Toolbox and develop business impact assessments. As well as the contribution to WP3, WP2 also provides business and economic analysis of survey results to WP4 to establish business use cases, users' needs, and an extensive and empirical business analysis of BDT projects and impacts.

1.3.1 Interdependencies with WP3

The main objective of WP3 is the specification, design and delivery of the DataBench Toolbox. The Toolbox aims to add value for the BDA adoption community by 1) enabling intelligent search and selection of existing Big Data benchmarks categorised with rich metadata; 2) automating the process of downloading, deploying and executing selected benchmarks; 3) providing ways to integrate the technical results of the execution to achieve a homogenised set of technical metrics; and 4) bridging the gap to visualise business insights taking into account both the technical metrics and other business and use-case-specific aspects.

It is this last objective that characterises the relationship between WP2 and WP3. WP3 is related to Big Data technical benchmarking and homogenisation of technical metrics coming from technical benchmarks. WP2 will provide a way of assessing and interpreting users' needs from a business perspective. The two results will be combined and delivered to the user in a common framework inside the Toolbox user interface.

1.3.2 Interdependencies with WP4

The main goal of WP4 is to evaluate the business performance of BDT through case studies and the DataBench Toolbox. There is a tight and continuous interdependency between WP2 and WP4. At this stage in the project, WP2 survey data is available, as well as detailed results from the analysis of survey data conducted as part of WP2. This input provides the overall framework where WP4 case studies can be positioned and leveraged to gather a more reliable understanding of complex BDT-related phenomena. While WP2 survey represents an *extensive* empirical analysis of BDT projects, WP4 cases constitute an *in-depth* empirical investigation of the interrelated business and technical drivers of BDT projects. The complementarities between extensive and intensive research methods are well known and are generally considered extremely important for the scientific soundness of empirical research methodologies [1], [2]. Extensive research helps identify general high-level rules and in-depth analysis provides more articulate insights explaining how those rules can be interpreted and applied in real BDT use cases.

To identify the high-level rules driving the application of BDTs, the WP2 questionnaire has been analysed with both a qualitative and a more rigorous statistical approach. The general research focus of DataBench is on the relationship between business and technical choices

in BDT projects, to understand how technical benchmarks can help make the best technical choices and maximise business benefits. Survey data has been analysed from this perspective, to discover correlations between technical and business variables and to gather general insights on how companies have combined their business requirements with their technical choices in BDT projects. The general rules tying business and technical variables are summarised with a simple interpretive framework that is described in this report and represents a fundamental input for the in-depth case study analysis conducted in WP4 (preliminary insights from this analysis are reported in D4.2).[1]

1.4. Methodology Approach

1.4.1 Overview

As described in D.2.1 "Economic and Market Analysis Methodology", the methodology used in this report is based on the following steps implemented between May 2018 and January 2019.

Phase 1

a) Desk research of main public sources (mainly Eurostat and OECD) to select the most relevant economic indicators
b) Extraction of relevant data from IDC databases and ongoing research on BDTs and the European data market
c) Elaboration of data to identify the most economically significant industries and those with the highest potential Big Data impact
d) Preliminary classification of main use cases by industry and business process and main KPIs based on desk and IDC research
e) Primary data collection, through a survey of a casual sample of 700 European BDT business users, representative of the EU industry
f) Elaboration of the survey results on KPIs, BDTs used and use cases by industry, company size and country
g) Calculation of preliminary benchmarks of economic and industrial significance
h) Production of deliverable D.2.2 presenting the main results of the survey and the preliminary benchmarks

Phase 2

In this report we implemented the following steps:

1. In-depth elaboration of survey results:
    a. From a business point of view, KPI relevance is evaluated firstly by industry and then by company size. This is compared against use case relevance against both criteria.
    b. From a technical point of view, a descriptive and statistical analysis is performed on the technical questions of the survey.

1 Reference [1] Dubé, L., & Paré, G. (2003): Rigor in Information Systems Positivist Case Research: Current Practices, Trends, and Recommendations/MIS Quarterly, 27(4), 597-636. [2] Brannen, J. (Ed.). (2017): Mixing Methods: Qualitative and Quantitative Research/New York, NY: Routledge, 2017.

2. Drawing conclusions on the relevance and role of the industrial KPIs estimated in the first phase and on the main user needs of benchmarking by industry and by company size, with a specific focus on SMEs
3. Launching a new wave of invitations to the DataBench survey addressed to the industrial partners of the H2020 ICT projects 14 and 15 and others in the Big Data area to collect further data and evidence on KPIs
4. Producing this deliverable (D.2.3) focused on further analysis of users' needs

1.4.2 Survey of Business Users

This section summarises the DataBench survey methodology and approach. The survey was carried out in September–October 2018 among European business organisations in 11 Member States, resulting in 700 valid interviews segmented as follows:

- 11 Member States: France, Germany, the Netherlands, the UK, the Nordics (Denmark, Sweden), Southern Europe (Italy, Spain), CEE (Czech Republic, Poland, Romania)
- 16 industry sectors and 7 employment size classes (for more details see the Annex).

The survey excluded micro-enterprises with fewer than 10 employees (unlikely to be advanced adopters of BDT). The survey was conducted in local language by experienced interviewers, targeted senior decision makers and influencers for BDTs, and screened respondents on the basis of their actual or planned use of BDA. Business organisations not using and not interested in using BDTs were excluded.

The industry classification is based on Eurostat's NACE REV. 2 code to be able to use statistical data on value added and other parameters as well as IDC's Vertical Market databases. The following industries were excluded for the following reasons:

- Government: DataBench is focused on the private sector, government does not use the same business KPIs as the private sector, and the number of government agencies varies substantially from country to country so that Eurostat does not provide comparable statistics by number of entities.
- Education: a mostly public and no profit sector, very different from private industry, with vastly different dynamics of technology adoption by segment (primary school versus research and university, for example). Investigating it would have required a different type of survey and questionnaire.
- To achieve a reasonable sample size by industry we had to eliminate another industry. We chose construction, which, according to the EDM monitoring tool statistics, is a low user of BDTs, is highly fragmented and would have required greater screening efforts to identify data user companies.

The survey aimed to collect quantitative evidence on the BDT use cases prioritised in each industry, actual and planned, the KPIs used, why they are used, the potential impacts on business processes and their relevance for business strategies and objectives. The Annex includes further data on the survey respondents and the main use cases of BDTs.

2. Users' Needs by Industry Sector

2.1. Overview

The DataBench survey provides insight into how European companies match their business needs with the realities of implementing Big Data analytics. This chapter analyses the stated importance by respondents of the seven key business KPIs used in the survey to investigate how companies adopting BDTs are benchmarking, or intend to benchmark, the business impact of their adoption.

Question 4 of the survey asked respondents, all of whom were either currently using BDA, piloting or implementing it, or considering or evaluating it for future use: "How important are the following business key performance indicators for measuring the impact of your organisation's Big Data and analytics efforts?" Respondents were asked about the seven main KPI categories selected by the DataBench conceptual framework to measure the most relevant business impacts:

- Cost reduction
- Time efficiency
- Product/service quality
- Revenue growth
- Customer satisfaction
- Business model innovation
- Increase in the number of new products or services launched

The following chapter analyses how the KPIs that were selected as particularly relevant to benchmarking the impact of BDA initiatives by the respondents varied by industry sector and company size.

Big Data analytics has the potential to be a key component of digital transformation. A formal assumption of IDC's Big Data and Analytics Software Market Forecasting[2] is that "digital transformation (DX) continues to be a key driver for growth in BDA software spending. DX initiatives, driven from top-level executives, are resulting in a deep assessment of current business practices and demands for better, faster and more comprehensive access to data and related analytics and insights." However, the necessary level of investment and disruptive change to IT infrastructure and the associated business processes to achieve the anticipated business benefits is considerable. Most companies, therefore, adopt either an incremental approach (via proof of concept) and/or a focused approach (addressing use cases in one or more business areas, justified by anticipated KPI improvements).

Priorities for identifying where to initially invest in BDA projects will inevitably vary from company to company.

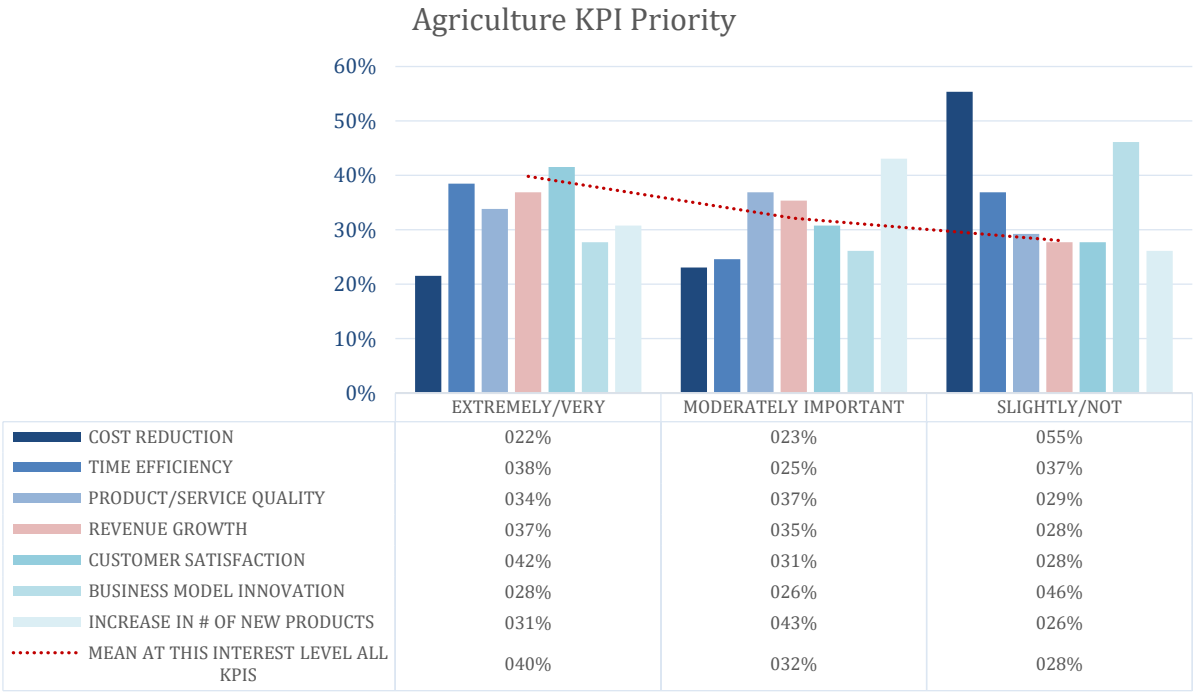
2 Source: IDC Worldwide Big Data and Analytics Software Forecast, 2018–2022 (IDC #US44243318, www.idc.com)

These priorities are influenced by factors such as:

- Industry sector
- Company size
- Existing BDA-relevant capability; some business units may have significant existing BDA capability, in ICT infrastructure and/or business readiness, reducing the level of investment required for proof-of-concept projects in these areas, even if they are not necessarily where the maximum business benefits can be achieved

2.2. Anticipated KPI Contributions from BDA Across All Industry Sectors and Sizes

2.2.1 Importance of KPIs in the Agriculture Sector



	EXTREMELY/VERY	MODERATELY IMPORTANT	SLIGHTLY/NOT
COST REDUCTION	022%	023%	055%
TIME EFFICIENCY	038%	025%	037%
PRODUCT/SERVICE QUALITY	034%	037%	029%
REVENUE GROWTH	037%	035%	028%
CUSTOMER SATISFACTION	042%	031%	028%
BUSINESS MODEL INNOVATION	028%	026%	046%
INCREASE IN # OF NEW PRODUCTS	031%	043%	026%
MEAN AT THIS INTEREST LEVEL ALL KPIS	040%	032%	028%

Figure 4 — Importance of KPIs in the Agriculture Sector (% of respondents)

Source: DataBench Survey, October 2018, 700 interviews, industry sample size = 65

Agriculture Sector Higher Priorities

No KPIs are significantly more regarded as "extremely or very important" than in other sectors.

Agriculture Sector Lower Priorities

For the agriculture sector, significantly more survey respondents cited cost reduction, time efficiency, product/service quality, revenue growth, customer satisfaction and business model innovation as "slightly important or not important" goals for BDA adoption than in other sectors (with cost reduction especially so). Product/service quality and customer satisfaction are also significantly less often cited as a high priority.

Agriculture Sector Analysis

The survey results indicate that the expectations for KPI improvements in the sector are more modest than in other sectors.

The potential impact of BDA in the agricultural sector is impeded to some extent by the inevitable inflexibility of core production processes and the historical trend of slow investment in information technology.

Future productivity breakthroughs are still expected to lie mainly in R&D-led improvements such as gene editing, cellular agriculture or biofuel crops.

However, there are numerous opportunities for BDA to provide positive business improvements in the sector and these are being actively pursued, with venture investment in agricultural technology rising quickly from an initially low level compared with other sectors.

A key opportunity is precision agriculture, which involves collecting and analysing data about production down to potentially the individual plant level to increase productivity and ensure that best practice policies are adhered to. Key technologies are remote sensing via satellite and drone observation and on-the-spot IoT sensors. In each case, analytics are key to the interpretation of observations and recommendations for appropriate action.

Other BDA application areas include optimisation and sharing of equipment use via digital marketplaces, predictive maintenance for expensive equipment items, and intelligent sharing of information on skills and trends between industry practitioners.

For example, the Horizon 2020 project DataBio is piloting 26 projects applying BDA in the sector with applications including local produce price prediction, fisheries planning and weather risk assessment, with many focusing on remote sensing to achieve precision agriculture.

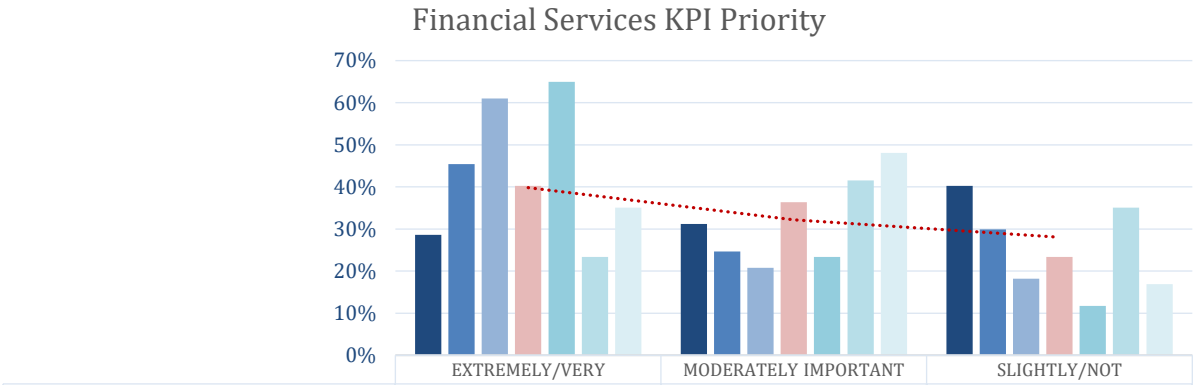
As suggested by the survey results, there do not seem to be as many BDA-related transformational opportunities in the sector, which is tied to its underlying biological processes (as with some other sectors). However, there are a range of important applications, centred around achieving basic productivity improvements. As agricultural IoT technology becomes more widespread, the importance of BDA in the sector will increase.

Comparison with Key Sector Use Cases

As noted in Annex 7.5, from the previous deliverable (D2.2), the top three use cases highlighted by respondents in the sector are field mapping and crop scouting, price optimisation, and inventory and service parts optimisation.

The first use case is related to the precision agriculture use case (which is also in the top 10 use cases for the sector) and an innovative application of BDA; the latter two are illustrative of less transformative applications of BDA in the sector, but which still have the potential to improve efficiency.

2.2.2 Importance of KPIs in the Financial Services Sector



	EXTREMELY/VERY	MODERATELY IMPORTANT	SLIGHTLY/NOT
COST REDUCTION	029%	031%	040%
TIME EFFICIENCY	045%	025%	030%
PRODUCT/SERVICE QUALITY	061%	021%	018%
REVENUE GROWTH	040%	036%	023%
CUSTOMER SATISFACTION	065%	023%	012%
BUSINESS MODEL INNOVATION	023%	042%	035%
INCREASE IN # OF NEW PRODUCTS	035%	048%	017%
MEAN AT THIS INTEREST LEVEL ALL KPIS	040%	032%	028%

Figure 5 — Importance of KPIs in the Financial Services Sector (% of respondents)

Source: DataBench Survey, October 2018, 700 interviews, industry sample size = 77 interviews

Financial Services Sector Higher Priorities

Significantly more respondents see customer satisfaction as "extremely or very important" than in other sectors; significantly fewer respondents see it as "slightly important or not important".

Significantly fewer respondents in the financial services sector see cost reduction, customer satisfaction and an increase in the number of new products or services launched as "slightly important or not important" when it comes to goals for BDA adoption than in other sectors.

Financial Services Sector Lower Priorities

There are no KPIs in this sector that are considered significantly less important than in other sectors.

Financial Services Sector Analysis

The survey results indicate that customer satisfaction is the main driving force behind BDA adoption in financial services, with cost reduction and an increase in the number of new products or services also being rated slightly higher than in other sectors.

This is not entirely surprising, as the financial services industry has always been a leader in technical innovation, and customer satisfaction is the competitive differentiator between service providers, whether they be high street banks, insurance companies or investment management providers.

There is a long history of applying statistical and machine learning techniques to predicting market and individual asset values, but with the increasing availability of data sources on customer and potential customer behaviour, combined with the increasing maturity and usability of BDA tools, it has become increasingly possible to deploy new techniques to

personalise services for customers. An important application is in the effective cross-selling and up-selling of services that accurately match customer needs.

There are also novel applications for both real-time analysis and real-time data, which is becoming increasingly straightforward to source, integrate and analyse. Examples of the former include real-time-tailored insurance policies based on up-to-date information about the client and the risks to be insured.

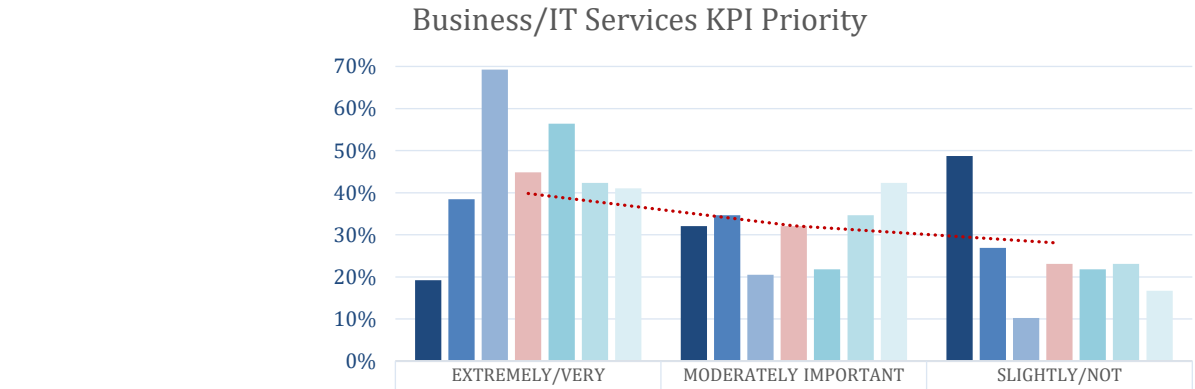
Examples of the latter include alerting insured drivers who are driving dangerously that this has been remotely observed. Such technology makes it possible to design new types of insurance or other financial services that are based on the availability of data to assess risks in real time or near real time.

Comparison with Key Sector Use Cases

The top three use cases highlighted by respondents in the sector are new product development (aligned with an increase in the number of new products or services launched); customer profiling, targeting and optimisation of offers (aligned with customer satisfaction); and risk exposure assessment (aligned with cost reduction).

There seems to be a good match between use cases where BDA is or expected to be deployed and the expected KPI benefits.

2.2.3 Importance of KPIs in the Business and IT Services Sector



	EXTREMELY/VERY	MODERATELY IMPORTANT	SLIGHTLY/NOT
COST REDUCTION	019%	032%	049%
TIME EFFICIENCY	038%	035%	027%
PRODUCT/SERVICE QUALITY	069%	021%	010%
REVENUE GROWTH	045%	032%	023%
CUSTOMER SATISFACTION	056%	022%	022%
BUSINESS MODEL INNOVATION	042%	035%	023%
INCREASE IN # OF NEW PRODUCTS	041%	042%	017%
MEAN AT THIS INTEREST LEVEL ALL KPIS	040%	032%	028%

Figure 6 — Importance of KPIs in the Business and IT Services Sector (% of respondents)

Source: DataBench Survey, October 2018, 700 interviews, industry sample size n = 78

Business and IT Services Sector Higher Priorities

Significantly more respondents see product/service quality as "extremely or very important" than in other sectors; it is also significantly less cited as "slightly important or not important".

Business model innovation is significantly more regarded as "extremely or very important" than in other sectors.

Significantly more respondents in this sector said an increase in the number of new products or services launched is "extremely or very important", while significantly fewer said it is "slightly important or not important".

Business and IT Services Sector Lower Priorities

There are no KPIs in this sector that are considered significantly less important than in other sectors.

Business and IT Services Sector Analysis

The survey results indicate that product/service quality and an increase in the number of new products or services launched are the main driving force behind BDA adoption in business and IT services, with business model innovation also important and other KPI drivers being comparable to other sectors.

It is significant that product/service quality is regarded as a more important goal for BDA deployment than customer satisfaction, as in the financial services sector. This suggests a slightly different emphasis in how companies regard their business model: the business and IT services sector concentrates on the quality of its offering with the expectation that improvements here will directly lead to increases in market share and/or customer satisfaction.

Suppliers of business and IT services have always faced resource management issues that parallel but are subtly different from those in other sectors such as manufacturing and retail:

- Forecasting demand and being able to allocate appropriate resources
- Planning the acquisition, training and management of human resources to a greater degree than other sectors
- Allocating resources to clients and other projects
- Competitive pricing of often completely bespoke service offerings

BDA can improve effectiveness in these areas, all of which impact most directly on the quality of services offered, in accordance with the findings of the survey.

There are also numerous ways that BDA can improve the quality of different types of service offering.

The scope for BDA to support the introduction of new types of products and services, the other distinctive KPI goal of this sector for BDA, can be divided into:

- Better analysis of client requirements and buying patterns. This parallels BDA use cases in most other sectors.
- New products and services based on BDA capabilities developed by the service provider and offered to clients. These are potentially more numerous and diverse than in any other sector due to service providers' flexibility when it comes to introducing new offerings.

Such benefits are as diverse as the sector, but include the following examples. In audit and management consulting, improvement of the speed and accuracy of internal and external audits can be achieved by automated recognition of risks and patterns within a client

organisation's activities. New service opportunities include the ability to anticipate future risks and make recommendations based on analytics, and the potential to offer new rapid and potentially real-time internal audit services.

Similar advantages can be achieved in the overlapping field of forensic data analytics, which also extends to information security and compliance management services.

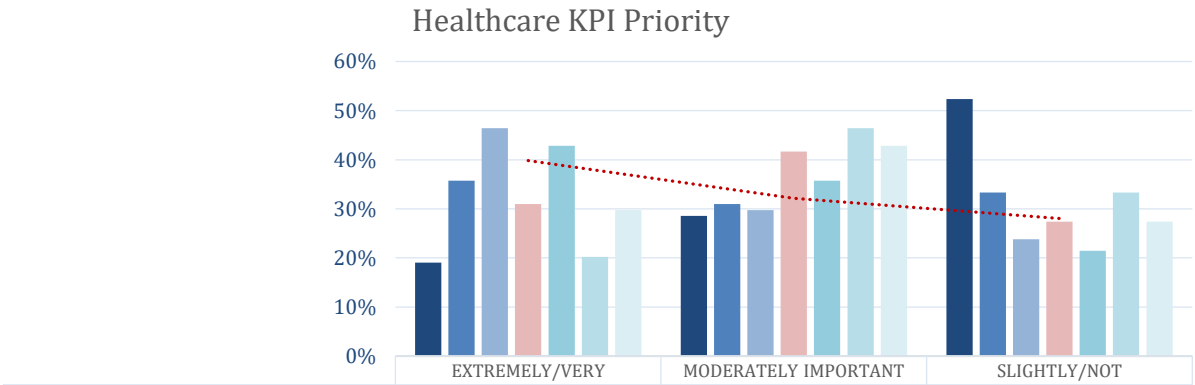
In IT service provision, BDA can not only streamline resource management within the supplier organisation, it can also enable the service provider to analyse the client's IT systems and operations before, during and after engagement.

In addition, and very importantly for the sector, consulting companies' perceived level of awareness of BDA use cases and supporting technologies, and how these can best address their clients' problems and improve their KPIs, is becoming a key selling proposition for their services.

Comparison with Key Sector Use Cases

Respondents cited a wide range of use cases, as would be expected in a diverse sector. Customer profiling, targeting and optimisation of offers is the most mentioned use case. However, risk exposure assessment, price optimisation and new product development are all at similar levels, suggesting there is a split in priorities between this kind of tactical deployment of BDTs and more transformational applications such as number of new products or services launched.

2.2.4 Importance of KPIs in the Healthcare Sector



	EXTREMELY/VERY	MODERATELY IMPORTANT	SLIGHTLY/NOT
COST REDUCTION	019%	029%	052%
TIME EFFICIENCY	036%	031%	033%
PRODUCT/SERVICE QUALITY	046%	030%	024%
REVENUE GROWTH	031%	042%	027%
CUSTOMER SATISFACTION	043%	036%	021%
BUSINESS MODEL INNOVATION	020%	046%	033%
INCREASE IN # OF NEW PRODUCTS	030%	043%	027%
MEAN AT THIS INTEREST LEVEL ALL KPIS	040%	032%	028%

Figure 7 — Importance of KPIs in the Healthcare Sector (% of respondents)

Source: DataBench Survey, October 2018, 700 interviews, industry sample size n = 84

Healthcare Sector Higher Priorities

There are no KPIs that are considered to be significantly higher priorities for BDA impact than in other sectors.

Healthcare Sector Lower Priorities

Time efficiency is significantly less regarded as "extremely or very important" than in other sectors.

Revenue growth is significantly less regarded as "extremely or very important" than in other sectors and is also significantly more regarded as "slightly important or not important".

Customer satisfaction is significantly less regarded as "extremely or very important" than in other sectors.

Business model innovation is also significantly more regarded as "extremely or very important" than in other sectors.

Healthcare Sector Analysis

There is a wide range of clearly beneficial applications for BDA in healthcare, and the expectations for KPI contributions from BDA are fairly balanced, with no KPI significantly more focused on than in other sectors.

For example, resource allocation (R&D resources, laboratory, diagnostic or treatment equipment, pre/during/post-treatment human resources and treatment accommodation, etc.) is especially complex in this sector and software analysis has been effectively applied here for several decades. The increasing availability of data and the computational resources to process data volume and complexity can potentially significantly improve resource usage.

Numerous other BDA applications are based on the analysis of medical data, both at a population and patient level, to predict the appropriate strategic or tactical responses.

Analysis of telemedicine results, again at both a general and individual patient level, is a new and growing opportunity and benefit to the sector. This potentially includes not only data from specialised IoT medical sensors, but also from sources such as medical practitioner consultation apps and exercise and lifestyle data voluntarily made available for medical interpretation by individuals (collected by smart watches or wearables, for example).

It is perhaps surprising that time efficiency was not rated as extremely or very important by the survey respondents. The healthcare sector has one of the most complex human resource allocation requirements of any sector, with numerous types of specialists needing to be scheduled to interact with patients or members of the public for proactive healthcare, and all at the right times. For those needing diagnosis or treatment, any delays can have critical consequences.

Revenue growth and customer satisfaction are also not regarded as important KPIs for BDA deployment, presumably because an increase in overall efficiency is considered the primary target, with customer satisfaction and, in subsectors where this is possible, revenue growth arising as a result rather than being the primary target.

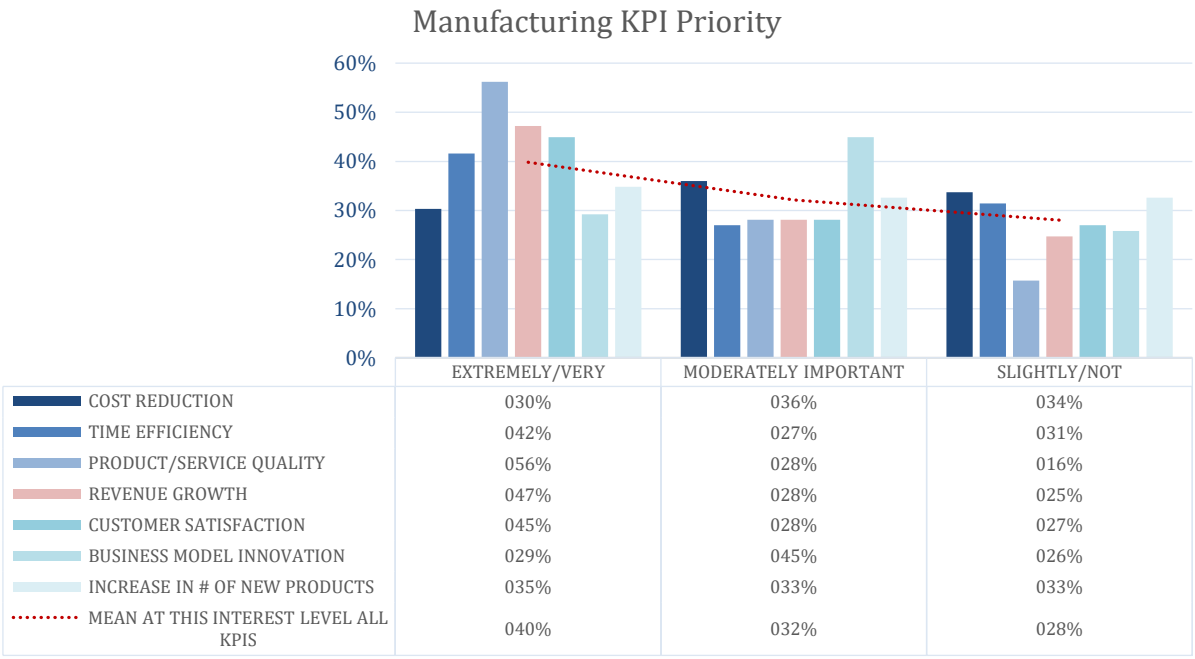
It may also be surprising that business model innovation is considered a lower priority than in other sectors. Worldwide concerns about the rising cost of healthcare combine with the increasing amount of data about how it is consumed to potentially enable the development of new models of funding and health insurance to address the challenges of changing demographics and healthcare needs. It is possible that providing evidence to support transformation in healthcare provisioning is considered beyond the scope of what can be achieved via analysis of data arising from existing systems.

Comparison with Key Sector Use Cases

The top three use cases highlighted by respondents in the sector are regulatory intelligence, fraud prevention and detection, and quality of care optimisation.

These align with the apparent goal of broad efficiency increases, as do three of the next four most often highlighted use cases, albeit with slightly lower levels of citation (44%–47%): automated customer service, risk exposure assessment, patient admission and re-admission reductions, and illness/disease diagnosis and progression. It appears that the respondents consider that availability of, and the ability to integrate and analyse, large volumes of operational data makes BDA most valuable in a range of six broadly operational use cases. The use of analytics directly for illness/disease diagnosis and progression, which sector outsiders might expect to be one of the most important use cases in the healthcare sector, was only actually ranked seventh in the survey.

2.2.5 Importance of KPIs in the Manufacturing Sector



	EXTREMELY/VERY	MODERATELY IMPORTANT	SLIGHTLY/NOT
COST REDUCTION	030%	036%	034%
TIME EFFICIENCY	042%	027%	031%
PRODUCT/SERVICE QUALITY	056%	028%	016%
REVENUE GROWTH	047%	028%	025%
CUSTOMER SATISFACTION	045%	028%	027%
BUSINESS MODEL INNOVATION	029%	045%	026%
INCREASE IN # OF NEW PRODUCTS	035%	033%	033%
MEAN AT THIS INTEREST LEVEL ALL KPIS	040%	032%	028%

Figure 8 — Importance of KPIs in the Manufacturing Sector (% of respondents)

Source: DataBench Survey, October 2018, 700 interviews, sample size n = 89

Manufacturing Sector Higher Priorities

Cost reduction is significantly less regarded as "slightly important or not important" as a goal for BDA adoption in the manufacturing sector when compared with other sectors.

Manufacturing Sector Lower Priorities

There are no KPIs in this sector that are significantly less important than in other sectors.

Manufacturing Sector Analysis

It is not surprising that the sector has a wide range of KPIs that are expected to be improved with BDA implementation.

There are numerous use cases for BDA in the sector, including:

- Predictive maintenance for machinery
- Optimisation of automated product testing via machine learning
- Real-time integration of supply chains, as in the retail sector but often with more complexity
- Resource optimisation and a reduction in waste

The opportunities to benefit from BDA in the sector have been boosted by:

- The ability to effectively integrate and analyse the significant amounts of data already available from existing sources, such as manufacturing system monitoring data and information from traditional IT systems such as inventory, production, sales and support.
- The increasing availability of new data. The most-often-cited example of new data availability is from industrial IoT, whether during the manufacturing process or during use of the resulting products. In B2C contexts there is also the potential to automatically collect and analyse consumer feedback.

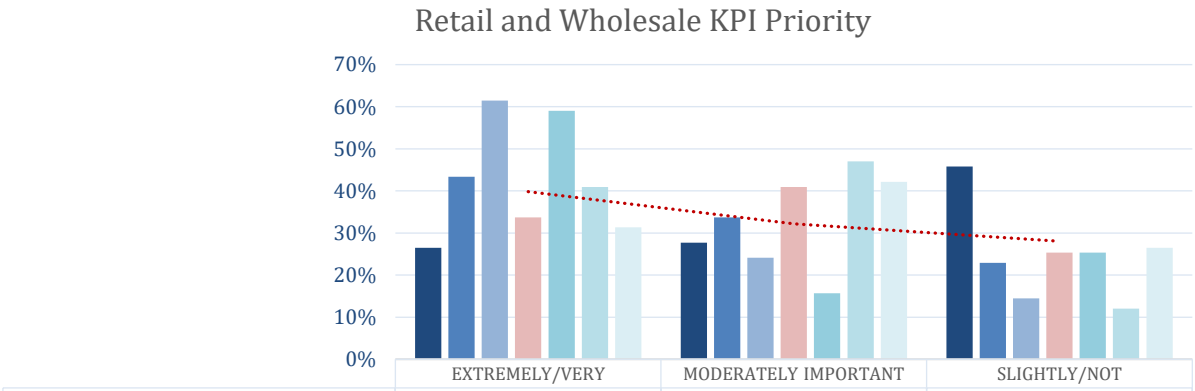
Not only is this new (or newly accessible) data now available for analysis, but it is increasingly feasible to apply advanced analytical techniques in a cost-effective way to target use cases that will deliver against the sector's KPIs.

Much of the manufacturing sector investment in BDTs has until recently been by the larger industry-leading manufacturers, but the benefits of BDA are rapidly becoming more accessible to midtier or smaller manufacturers.

Comparison with Key Sector Use Cases

The top three use cases highlighted by respondents in the sector are regulatory intelligence, new product development and price optimisation; the next-highest are supply chain optimisation and predictive maintenance. This is consistent with the wide range of potentially high-value BDA applications across the sector.

2.2.6 Importance of KPIs in the Retail and Wholesale Sector



	EXTREMELY/VERY	MODERATELY IMPORTANT	SLIGHTLY/NOT
COST REDUCTION	027%	028%	046%
TIME EFFICIENCY	043%	034%	023%
PRODUCT/SERVICE QUALITY	061%	024%	014%
REVENUE GROWTH	034%	041%	025%
CUSTOMER SATISFACTION	059%	016%	025%
BUSINESS MODEL INNOVATION	041%	047%	012%
INCREASE IN # OF NEW PRODUCTS	031%	042%	027%
MEAN AT THIS INTEREST LEVEL ALL KPIS	040%	032%	028%

Figure 9 — Importance of KPIs in the Retail and Wholesale Sector (% of respondents)

Source: DataBench Survey, October 2018, 700 interviews, industry sample size n = 83

Retail and Wholesale Sector Higher Priorities

Business model innovation is significantly more regarded as "extremely or very important" than in other sectors and is significantly less regarded as "slightly important or not important".

Time efficiency is significantly less regarded as "slightly important or not important" as a goal for BDA adoption when compared with other sectors.

Retail and Wholesale Sector Lower Priorities

Revenue growth is significantly less regarded as "extremely or very important" than in other sectors.

Retail and Wholesale Sector Analysis

For decades, the retail sector has intensively used analytics in applications including supply chain management, inventory control and stocking/pricing by analysing consumer behaviour via touch points such as discount offers and loyalty programmes.

Brick-and-mortar shopping is under extreme pressure from online retailing. Advanced BDA offers the potential to respond by optimising in-store product stocking, pricing and customer offers, especially where retailers can analyse both online and in-store customer behaviour.

The introduction of successful innovative business models in related sectors (for example, Uber and Airbnb) has added to the considerable interest in exploring options to change the retailer/customer relationship in this turbulent sector. Examples include novel product rental options (for example, for fashion clothing and high-value domestic appliances) and automatic replenishment of domestic consumables. Both these business models have been

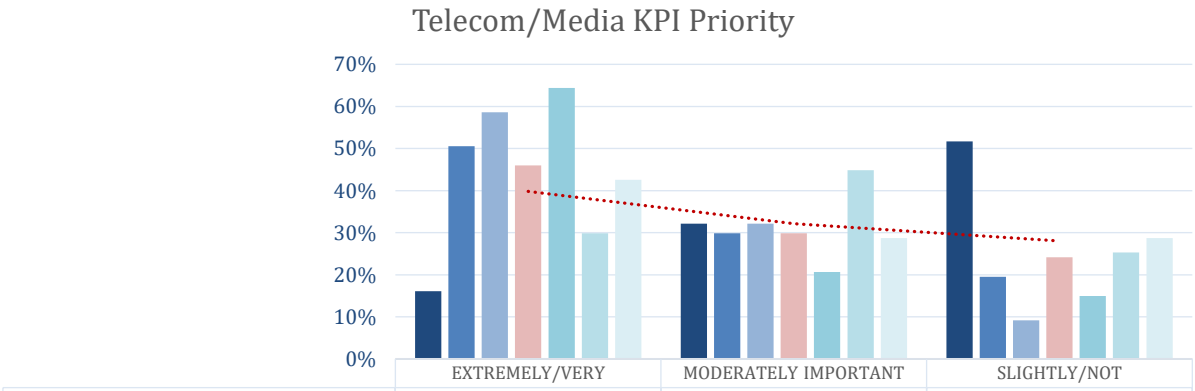
successfully applied in retail in the past, but accurate analytics are vital to inform organisations on whether they might now be viable options.

It is perhaps surprising that revenue growth is significantly less regarded as "extremely or very important", given the traditional preoccupations of the sector; it seems likely that the search for new business models and products is driven by the search for increased revenues and profits.

Comparison with Key Sector Use Cases

The top rated use cases highlighted by respondents in the sector are new product development, price optimisation and supply chain optimisation, with intelligent fulfilment in fourth. The first and fourth of these can be directly related to the highest-ranked KPI, business model innovation, and the third to time optimisation. Price optimisation does not appear to relate directly to the more favoured KPIs, but is a clear traditional application of BDA in the sector, for example in market basket analysis, going back to at least the 1990s.

2.2.7 Importance of KPIs in the Telecom/Media Sector



	EXTREMELY/VERY	MODERATELY IMPORTANT	SLIGHTLY/NOT
COST REDUCTION	016%	032%	052%
TIME EFFICIENCY	051%	030%	020%
PRODUCT/SERVICE QUALITY	059%	032%	009%
REVENUE GROWTH	046%	030%	024%
CUSTOMER SATISFACTION	064%	021%	015%
BUSINESS MODEL INNOVATION	030%	045%	025%
INCREASE IN # OF NEW PRODUCTS	043%	029%	029%
MEAN AT THIS INTEREST LEVEL ALL KPIS	040%	032%	028%

Figure 10 — Importance of KPIs in the Telecom and Media Sector (% of respondents)

Source: DataBench Survey, October 2018, 700 interviews, sample size n = 87

Telecom/Media Sector Higher Priorities

Time efficiency is significantly more regarded as "extremely or very important" than in other sectors and is significantly less regarded as "slightly important or not important".

Product/service quality is significantly less regarded as "slightly important or not important" as a goal for BDA adoption than in other sectors.

Customer satisfaction is also significantly more regarded as "extremely or very important" than in other sectors and is significantly less regarded as "slightly important or not important".

An increase in the number of new products or services launched is significantly more regarded as "extremely or very important" than in other sectors.

Telecom/Media Sector Lower Priorities

Cost reduction is significantly less regarded as "extremely or very important" than in other sectors.

Telecom/Media Sector Analysis

The telecoms sector faces an unprecedented challenge in consumer choice, as it is now easier than ever before for consumers to switch provider. Customer satisfaction is a critical factor in customer loyalty, especially when it comes to signal strength and availability of mobile devices.

Advanced analytics, especially in real time, can greatly increase the ability to respond more quickly, both to address technical service issues with the operator that may affect customer satisfaction and to rapidly and proactively interact with customers who have experienced issues that may affect their satisfaction.

For some time the telecoms sector has pioneered the use of analytics in customer loyalty assessment (customer churn prediction) and now has access to considerably more data about how customers use and experience their services. The availability of more powerful BDA can inform both strategic decisions (including investment in infrastructure) and tactical decisions about individual customers.

Improved BDA also enables more accurate analysis of the past, current and potential future behaviour of a provider's infrastructure, enabling expensive decisions about investment in improving product/service quality to be made with more confidence.

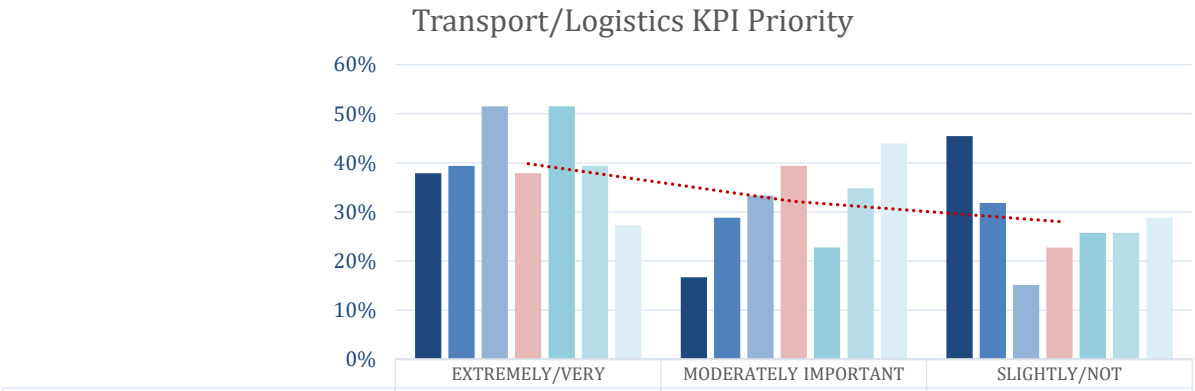
In the media sector, competition has also significantly increased, notably between content streaming services. Providers now have access to a wide range of data about current and potential future subscribers, and the capacity to analyse it. In many cases this now includes additional data such as not only what and when customers watch, but on which devices and potentially where. Planning of future offerings (an increase in the number of new products or services) can be informed by analysis of such data, as can effective personalisation (increasing customer satisfaction).

Comparison with Key Sector Use Cases

The two top use cases highlighted by respondents in the sector are customer profiling, targeting and optimisation of offers, and product and service recommendation systems. These are consistent with the favoured customer satisfaction and product/service quality KPIs.

The next most-cited use cases (which were fairly evenly mentioned) are price optimisation, regulatory intelligence and automated customer service. Apart from regulatory intelligence, these are also consistent with the customer satisfaction and product/service quality KPIs. Regulatory intelligence is a sector must-have capability and, like information security, has always been somewhat difficult to relate to "traditional" core KPIs.

2.2.8 Importance of KPIs in the Transport/Logistics Sector



	EXTREMELY/VERY	MODERATELY IMPORTANT	SLIGHTLY/NOT
COST REDUCTION	038%	017%	045%
TIME EFFICIENCY	039%	029%	032%
PRODUCT/SERVICE QUALITY	052%	033%	015%
REVENUE GROWTH	038%	039%	023%
CUSTOMER SATISFACTION	052%	023%	026%
BUSINESS MODEL INNOVATION	039%	035%	026%
INCREASE IN # OF NEW PRODUCTS	027%	044%	029%
MEAN AT THIS INTEREST LEVEL ALL KPIS	040%	032%	028%

Figure 11 — Importance of KPIs in the Transport and Logistics Sector (% of respondents)

Source: DataBench Survey, October 2018, 700 interviews, sample size n = 66

Transport/Logistics Sector Higher Priorities

Business model innovation is significantly more regarded as "extremely or very important" than in other sectors.

Cost reduction is significantly more regarded as "extremely or very important" than in other sectors on average. This is the only sector where this KPI was given this level of importance by a substantial proportion (38%) of respondents. However, 45% of respondents still rated this KPI as "slightly important or not important".

Transport/Logistics Sector Lower Priorities

An increase in the number of new products or services launched is significantly less regarded as "extremely or very important" as a goal for BDA adoption than in other sectors.

Transport/Logistics Sector Analysis

For many years companies in the sector have relied on software to ensure quality of service. The use of BDA to provide added value is now seen as a significant additional option.

Cost reduction is a core selling point in the sector, with timeliness and reliability regarded as a given.

Real-time tracking of deliveries from supplier to client (at least for the final delivery stage) has been both an operational management resource and a customer benefit for some time. Improved BDA technology enables this information to be analysed much more effectively for delivery optimisation, combined with basic data about the frequency and value of deliveries to regions and destinations.

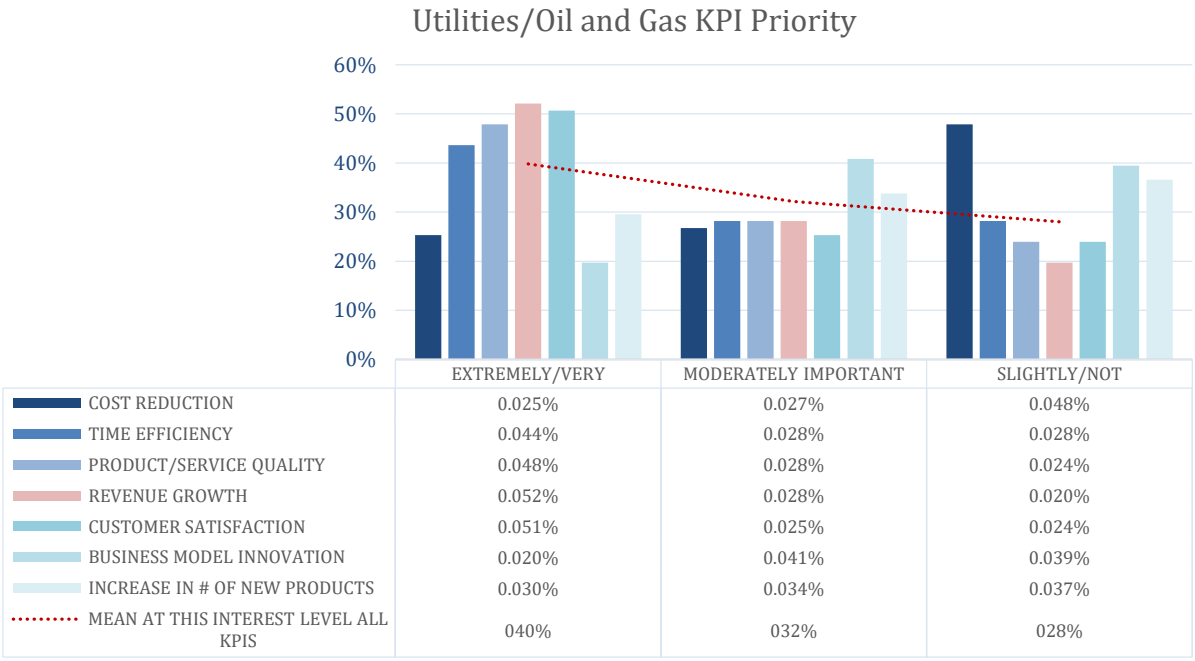
Additional competition in the sector has been fuelled by IT-based innovations such as the increased ability to outsource segments of the delivery chain in real time, including real-time optimisation of bulk/shared deliveries and returns, both in B2B and B2C contexts, and the potential for logistics suppliers to cooperate with their clients (even with their competitors) to ensure the most effective delivery of goods.

Accordingly, business model innovation is considered a significant objective for BDA investment in the sector. Such process changes are both expensive and risky, and such decisions need to be based on thorough and reliable analysis.

Comparison with Key Sector Use Cases

The three top use cases highlighted by respondents in the sector are logistics and package delivery management, new product development, and inventory and service parts optimisation. These are completely aligned with the top-ranked KPIs for the sector.

2.2.9 Importance of KPIs in the Utilities/Oil and Gas Sector



	EXTREMELY/VERY	MODERATELY IMPORTANT	SLIGHTLY/NOT
COST REDUCTION	0.025%	0.027%	0.048%
TIME EFFICIENCY	0.044%	0.028%	0.028%
PRODUCT/SERVICE QUALITY	0.048%	0.028%	0.024%
REVENUE GROWTH	0.052%	0.028%	0.020%
CUSTOMER SATISFACTION	0.051%	0.025%	0.024%
BUSINESS MODEL INNOVATION	0.020%	0.041%	0.039%
INCREASE IN # OF NEW PRODUCTS	0.030%	0.034%	0.037%
MEAN AT THIS INTEREST LEVEL ALL KPIS	040%	032%	028%

Figure 12 — Importance of KPIs in the Utilities/Oil and Gas Sector (% of respondents)

Source: DataBench Survey, October 2018, 700 interviews, industry sample size n = 71

Utilities/Oil and Gas Sector Higher Priorities

Revenue growth is significantly more regarded as "extremely or very important" than in other sectors and is significantly less regarded as "slightly important or not important".

Utilities/Oil and Gas Sector Lower Priorities

Business model innovation is significantly more regarded as "slightly important or not important" than in other sectors and is significantly less regarded as "extremely or very important".

An increase in the number of new products or services launched is significantly more regarded as "slightly important or not important" than in other sectors.

Utilities/Oil and Gas Sector Analysis

The petrochemical sector has pioneered the use of advanced analytics in resource exploration, and more advanced BDA techniques will be deployed in this area, where exploration companies are well resourced to take advantage of new technologies.

Advanced BDA is also very widely used in downstream processing. McKinsey estimated in 2017 that "oil and gas has a $200 billion [processing] performance gap" which can be addressed by using advanced analytics.

On the supply side, utility company competition has been widely encouraged, leading to greater interest in BDA use cases ranging from customer relationship improvement (for example, churn analysis) to cost reduction via analysis of the greatly increasing volume of analysable data from smart metres. Potentially, such analysis could be extremely valuable in demand prediction (where the sector has been deploying machine-learning technology for some time based on the limited information from locally pooled metering).

It is perhaps surprising that business model innovation is not regarded as a BDA priority in the sector, given well documented opportunities such as the sharing/resale of consumer-generated energy. However, such innovations have tended to be hampered by technology cost and complexity rather than lack of analysis. When such non-traditional energy distribution models become more widespread, BDTs will be vital to ensuring that they operate effectively.

Comparison with Key Sector Use Cases

The three top use cases highlighted by respondents in the sector are predictive maintenance, field service optimisation and regulatory intelligence. These are clearly important for the cost reduction KPI, but not obviously directly related to the revenue growth KPI cited as the most important BDA driver for the sector. However, in many parts of the sector, cost is very much a primary driver for market share, so it seems likely that cost reduction is regarded as a proxy target to boost revenues.

2.3. **Concluding Remarks**

This chapter has analysed the survey results on the relative importance by industry of business KPIs used for benchmarking of BDTs, as presented in the previous D.2.2 deliverable (Figure 13), to investigate users' main needs as expressed in terms of their priorities. We have particularly focused on the less expected results, including traditionally important core KPIs such as cost reduction and time efficiency — which are now considered to be of low importance in almost all sectors and company sizes.

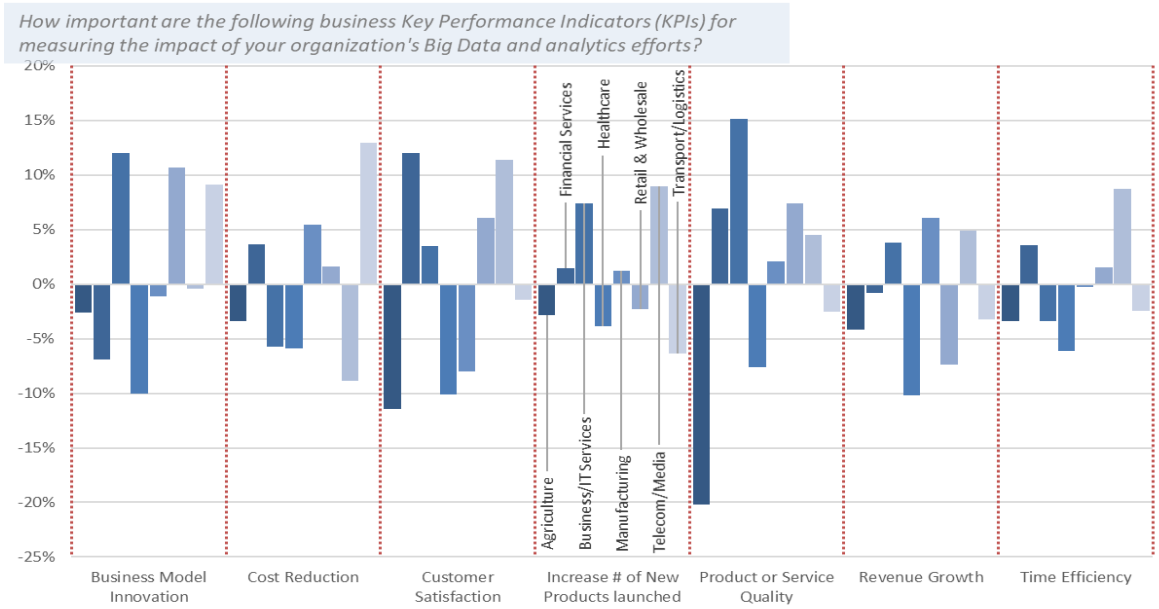


Figure 13 — KPI Ranking by Importance by Industry (% of respondents)

Source: IDC Elaboration of DataBench Survey, October 2018

We have identified outliers in the levels of interest in particular KPIs for each industry sector, focusing on KPIs where the level of interest at "very or extremely important" or "not at all or slightly important" was more than one standard deviation from the norm for that level of interest across all sectors. Figure 13 indicates the mean values across all respondents for each KPI priority level, against which sector mean preferences were compared.

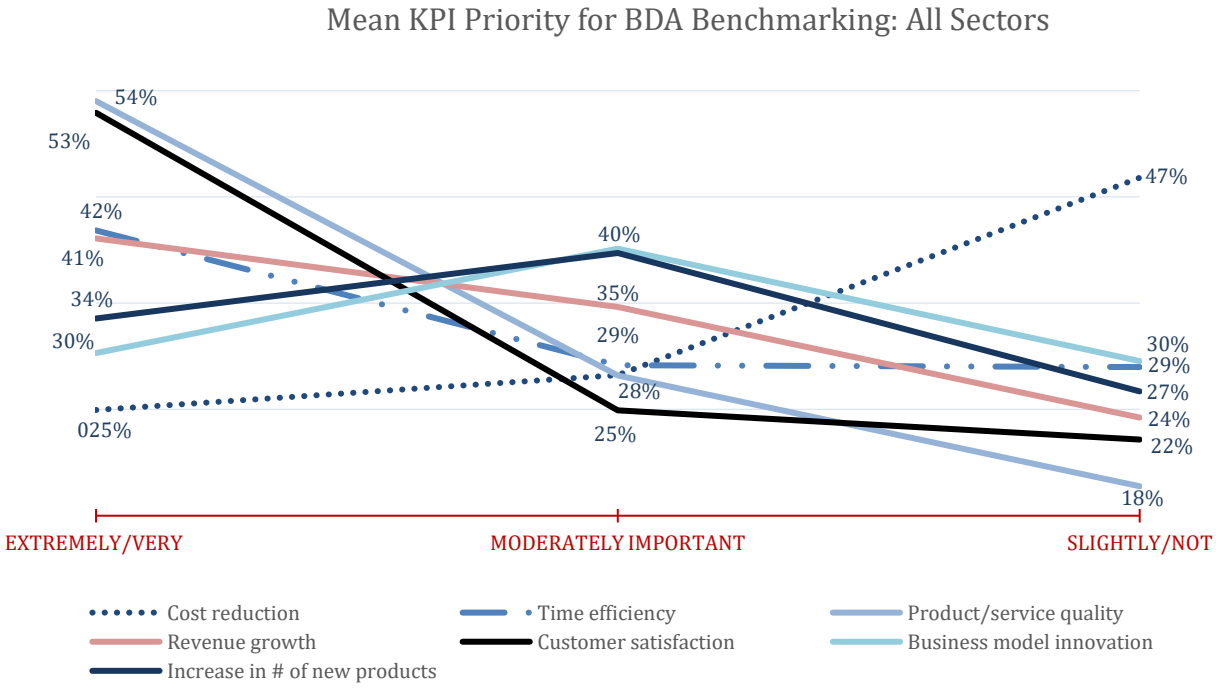


Figure 14 — Mean KPI Priority for BDA Benchmarking: All Sectors (% of respondents)

Source: DataBench Survey, October 2018, 700 interviews

Notable features include:

- Few respondents, on average, viewed KPIs negatively in terms of being "slightly important or not important". Average responses at this level are clustered between 18% and 30%, with only cost reduction averaging over 30%. This indicates an overall level of optimism that investment in BDA will bring a wide-ranging business impact.
- Overall, product/service quality and customer satisfaction are far more often rated as "very or extremely important", at over 50% on average, with time efficiency and revenue growth clustered in second place. This is consistent with demonstrated effective applications of BDA in customer relationship management, marketing, sales and support, and the expectation of BDA users and adopters that they can also implement such improvements.[3]
- KPI importance levels are generally quite consistent across all sectors, with exceptions noted previously. As business drivers in terms of KPI metrics do vary by sector, an understanding of these factors is of value to both users and adopters of BDTs and to vendors of products and services.

2.3.1 Considerations by Industry

In brief, the approaches to perceived benefits from BDA adoption vary as follows.

Agriculture

Unambitious across the board in KPI aspirations compared with all other sectors. However, the most-cited BDA use case is leading edge in terms of technology application (field mapping and crop scouting), indicating an appreciation of the potential for transformative BDA in the sector based on smart farming, in which BDA is a key component.

Financial Services

A strong focus on the well-established BDA impact on customer satisfaction, with cost reduction more regarded as a potential area of impact than in other sectors. It is also recognised that there is a potentially more transformative impact from an increase in the number of new products or services launched. There is a good match with mentioned use cases, as would be expected in a sector with a mature track record of BDA application.

Business/IT Services

A balance in expectations between potentially transformative (an increase in the number of new products or services launched, business model innovation) and incremental (product/service quality), with the most-cited use cases split accordingly.

Healthcare

A balanced expectation in terms of KPI impact, but with no KPIs regarded as significantly more important than in other sectors. Time efficiency, customer satisfaction and, especially, revenue growth are significantly less regarded as being important.

3 This is consistent with IDC research: the most important business goals driving BDA adoption are mostly customer related — optimisation of pricing strategies (41%), product service and programme improvement (42%) and improvements in customer understanding (43%) (European Businesses' Approach to Big Data Storage and Management, IDC #EMEA44911019, March 2019)

The top three most-mentioned use cases are, however, generally concerned with cost reduction. Given the number of potential BDA applications in the sector whose primary benefit is efficiency improvement, based on the large volume of data now available for analysis, this is indicative of the pragmatic approach of the sector to BDA implementation.

Manufacturing

Very broad expectations in terms of KPI impact from BDA, with a slightly higher emphasis on the potential for cost reduction than in other sectors (though still low). The most-often-cited use cases are consistent with these wide-ranging goals.

Retail and Wholesale

A balanced expectation in terms of KPI impact, including both potentially transformative business model innovation and operational improvement via time efficiency; somewhat comparable to the healthcare sector. The most-cited use cases are consistent with these.

Telecom/Media

Another balanced expectation, with an increase in the number of new products or services being potentially transformative, but also above-average expectations for time efficiency and customer satisfaction. The top-rated use cases are consistent with these expectations.

Transport/Logistics

Business model innovation has the highest level of expectation, but this is one of two sectors (the other being manufacturing) to rate cost reduction the highest. The top use cases are well aligned with these goals.

Utilities/Oil and Gas

This is the most conservative sector in terms of expectations of transformation from BDA, with revenue growth significantly more sought as a BDA outcome than any other sector, and business model innovation and an increase in the number of new products or services launched considered significantly less important. The most-cited use cases are concerned with efficiency and cost savings, which are consistent with aspirations for revenue growth in a sector where market share depends to a large extent on consumer pricing.

2.3.2 Sector Comparison by KPI Priority

This section summarises the distinctive attitudes to KPI importance described in the previous sections, measured by the indicator on the relative importance of the 7 main business KPIs by industry discussed in the previous sections. A distinctive attitude is determined by the sector mean response being above or below a standard deviation from the overall survey response. As shown in the "heat map" below, the attitudes towards the different KPIs range from very or moderately positive (hot: red/orange) to very or moderately negative (cool: light/dark blue). More specifically:

- Very positive is where the sector respondents distinctively considered the KPI to be both more "extremely/very more important" and also less "slightly/not less important" than for all respondents.
- Very negative is where the sector respondents distinctively considered the KPI to be both less "extremely/very more important" and also more "slightly/not less important" than for all respondents.

- Moderately positive is where the sector respondents either considered the KPI to be more "extremely/very more important" or less "slightly/not less important" than for all respondents.
- Moderately negative is where the sector respondents either considered the KPI to be less "extremely/very more important" or more "slightly/not less important" than for all respondents.
- Blank cells represent KPIs on which sector respondents present no relevant difference from all respondents and are therefore neutral.

	Cost Reduction	Time Efficiency	Product/ Service Quality	Revenue Growth	Customer Satisfaction	Business Model Innovation	Increase in # of New Products Launched
Agriculture							
Financial services							
Business/IT services							
Healthcare							
Manufacturing							
Retail and wholesale							
Telecom/media							
Transport/logistics							
Utilities/oil and gas							

Table 1 — Heat Map of Business KPI Priorities by Industry

The colour map can be read as follows:

- *Bright red: very positive*
- *Orange: moderately positive*
- *Dark blue: very negative*
- *Light blue: moderately negative*
- *Blank cell: neutral, i.e., consistent with the considered importance across all sectors*

Source: IDC Elaboration of DataBench Survey, October 2018

The sectors can be informally grouped into:

1. **Sectors that are conservative about the projected impact of BDA on KPIs.**

Agriculture and healthcare both fall into this category.

These are both sectors in which the application of IT has had generally patchy success, for a variety of reasons. In healthcare, regulation of the use and sharing of patient data is also considered to be an inhibitor to the widespread adoption of BDA.

As noted above, however, the most reported use case for agriculture is in the potentially transformative field of precision agriculture, which could enable the sector to achieve transformational effects via a combination of new technologies that have not previously been realistic.

This is consistent with the previous D2.2 DataBench deliverable (Preliminary Benchmarks of Industrial Significance of Big Data Technology Performance Parameters).

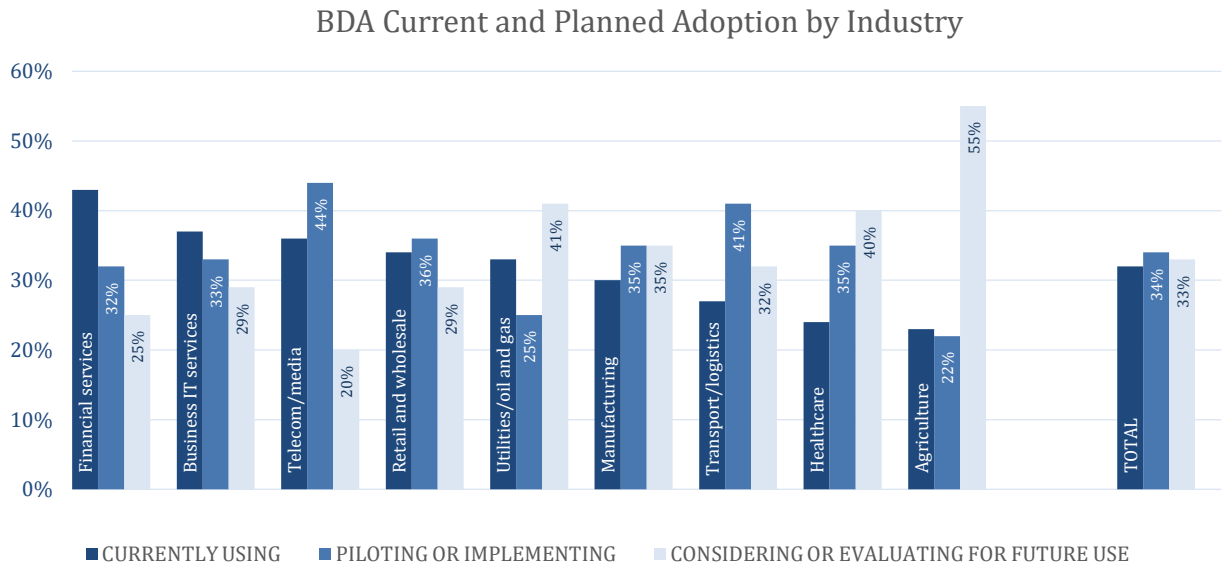


Figure 15 — Current and Planned BDA Adoption by Industry (% of respondents)

Source: DataBench deliverable D2.2

Figure 15 shows that healthcare and agriculture have the lowest levels of adoption. However, both industries demonstrate a strong intention to consider Big Data for future use.

2. **Sectors that are generally conservative about the projected impact of BDA on KPIs but which do anticipate potential transformational effects.**

The most clearly transformational KPIs are business model innovation and increase in the number of new products or services.

Retail and wholesale and transport/logistics have modest expectations in terms of BDA's contribution towards more operational business KPIs compared with other sectors, but both have high expectations for transformative KPIs.

Both sectors have benefited greatly from the application of IT, including traditional business intelligence, to increasing operational efficiency, but now anticipate more transformational impacts from the new generation of BDTs.

3. **Sectors with clear main-focus KPIs.**

Financial services and utilities/oil and gas both have clearly preferred KPIs for BDA (customer satisfaction and revenue growth respectively) but do not have clear expectations for transformative KPIs. Financial services foresee generally positive KPI contributions to business, however, as might be expected from its historical success in deploying analytics in a wide range of use cases.

4. **Sectors with Broad positive expectations, including transformation.**

The manufacturing sector has expectations for KPI delivery across all the surveyed KPIs that are in line with the overall average. This is consistent with the broad range of existing successful BDA use cases in the industry and the enormous potential for additional analysis from the range of technologies labelled Industry 4.0.

5. **Sectors aiming for both operational focus and transformational KPIs.**

Business/IT services and telecom/media have very clear KPI expectations for BDA in more traditional operational areas (product/service quality and time efficiency plus customer satisfaction respectively) but also in the increase in the number of new products or services and business model innovation respectively. This is not entirely surprising, given that IT services and telecom/media companies are very technology-literate and are in a good position to appreciate and implement the potential of modern BDTs.

	Category
Agriculture	Conservative
Healthcare	Conservative
Retail and wholesale	Conservative; seeking transformation
Transport/logistics	Conservative; seeking transformation
Manufacturing	Broad positive expectations, including transformation
Financial services	Clear main focus
Utilities/oil and gas	Clear main focus
Business/IT services	Aspirational for both operational focus and transformational KPIs
Telecom/media	Aspirational for both operational focus and transformational KPIs

Table 2 — Categorisation of Distinctive KPI Priorities by Sector

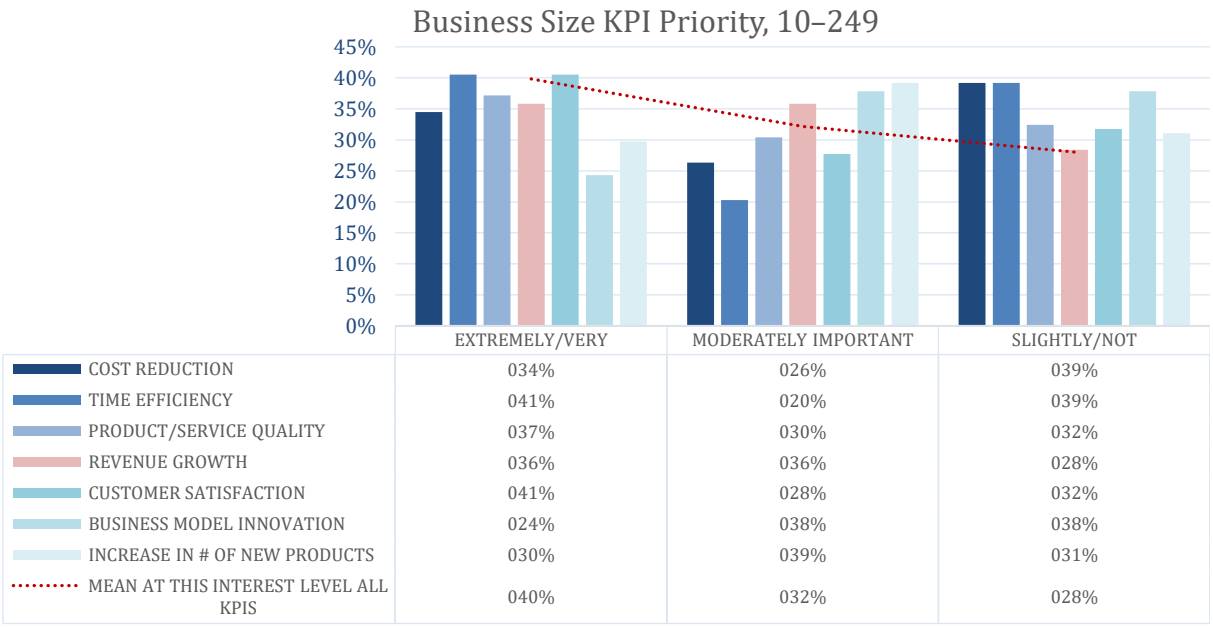
Source: IDC Elaboration of DataBench Survey, October 2018

3. Users' Needs by Organisation Size

As noted in the previous D2.2 DataBench deliverable (Preliminary Benchmarks of Industrial Significance of Big Data Technology Performance Parameters) there is a complex relationship between organisations' goals for BDA adoption, organisation size and the sector in which they operate. For example, in the manufacturing sector larger organisations have tended to be significantly more ambitious in the adoption of BDA. In other sectors, such as healthcare, the earliest adopters of BDA technology have often been small companies seeking to provide distinctive services to more conservative providers.

However, there are significant general trends, as illustrated by the survey results and analysed below.

3.1. Importance of KPIs in Companies with 10–249 Employees



	EXTREMELY/VERY	MODERATELY IMPORTANT	SLIGHTLY/NOT
COST REDUCTION	034%	026%	039%
TIME EFFICIENCY	041%	020%	039%
PRODUCT/SERVICE QUALITY	037%	030%	032%
REVENUE GROWTH	036%	036%	028%
CUSTOMER SATISFACTION	041%	028%	032%
BUSINESS MODEL INNOVATION	024%	038%	038%
INCREASE IN # OF NEW PRODUCTS	030%	039%	031%
MEAN AT THIS INTEREST LEVEL ALL KPIS	040%	032%	028%

Figure 16 — Importance of KPIs in SMEs (10–249 Employees)

Source: DataBench Survey, October 2018, 700 interviews, sample size n = 148

Companies with 10–249 Employees Higher Priorities

Cost reduction is significantly more regarded as "extremely or very important" than in other organisation size segments and is significantly less regarded as "slightly important or not important".

Companies with 10–249 Employees Lower Priorities

- Time efficiency is significantly more regarded as "slightly important or not important" as a goal for BDA adoption than in other organisation size segments.
- Product/service quality is significantly more regarded as "slightly important or not important" as a goal for BDA adoption than in other organisation size segments and less regarded as "extremely or very important".
- Revenue growth is significantly more regarded as "slightly important or not important" as a goal for BDA adoption than in other organisation size segments and less regarded as "extremely or very important".
- Customer satisfaction is significantly more regarded as "slightly important or not important" as a goal for BDA adoption than in other organisation size segments and less regarded as "extremely or very important".
- Business model innovation is significantly more regarded as "slightly important or not important" as a goal for BDA adoption than in other organisation size segments and less regarded as "extremely or very important".

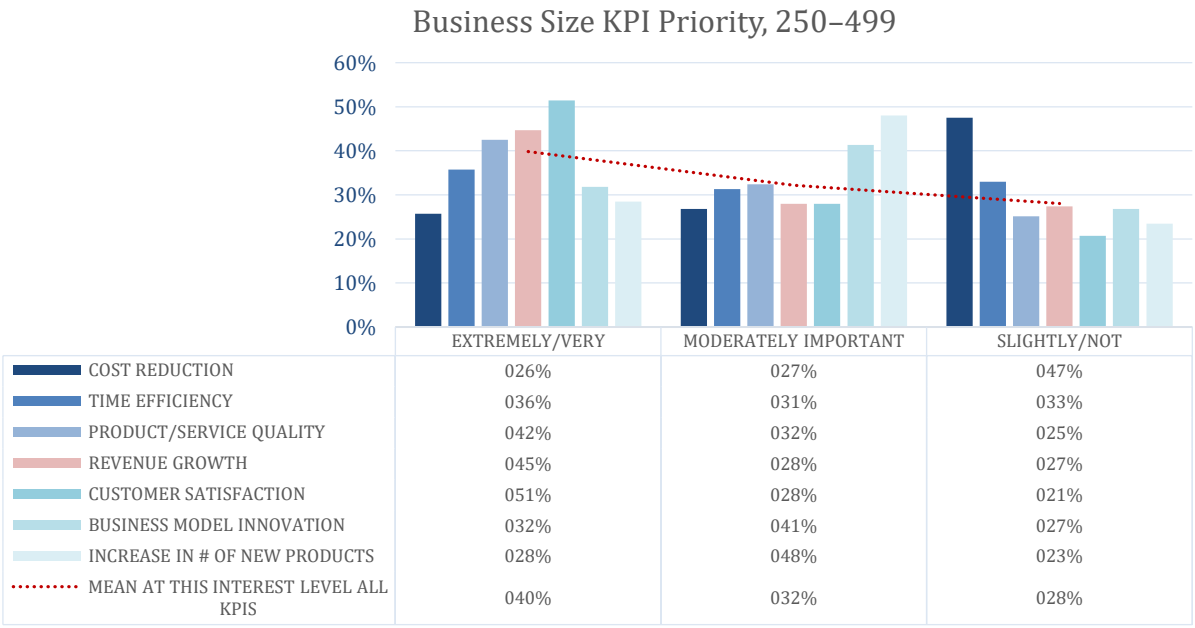
Companies with 10–249 Employees Segment Analysis

Companies that are pre-trading or only have a few customers are already committed to their business proposition and there is seldom the opportunity, the time or the funding to streamline delivery.

Within this size band, however, there are also established, traditional and long-standing companies, sometimes inherited as a family business, which do not generally aim to change the status quo in which they operate or seek very significant growth. Business model innovation is not usually a significant priority here either.

However, as the survey confirms, cost reduction and time efficiency are priorities for both categories of small company when they use or plan to use BDA.

3.2. Importance of KPIs in Companies with 250–499 Employees



	EXTREMELY/VERY	MODERATELY IMPORTANT	SLIGHTLY/NOT
COST REDUCTION	026%	027%	047%
TIME EFFICIENCY	036%	031%	033%
PRODUCT/SERVICE QUALITY	042%	032%	025%
REVENUE GROWTH	045%	028%	027%
CUSTOMER SATISFACTION	051%	028%	021%
BUSINESS MODEL INNOVATION	032%	041%	027%
INCREASE IN # OF NEW PRODUCTS	028%	048%	023%
MEAN AT THIS INTEREST LEVEL ALL KPIS	040%	032%	028%

Figure 17 — Importance of KPIs in Medium-Sized Enterprises (250–499 Employees)

Source: DataBench Survey, October 2018, 700 interviews, sample size n = 179

Companies with 250–499 Employees Higher Priorities

Revenue growth is significantly more regarded as "extremely or very important" than for other organisation size segments.

For the importance of "increase in the number of new products or services launched", there is a significant response for "moderately important", with fewer respondents opting for "extremely or very important" or "slightly important or not important".

Companies with 250–499 Employees Lower Priorities

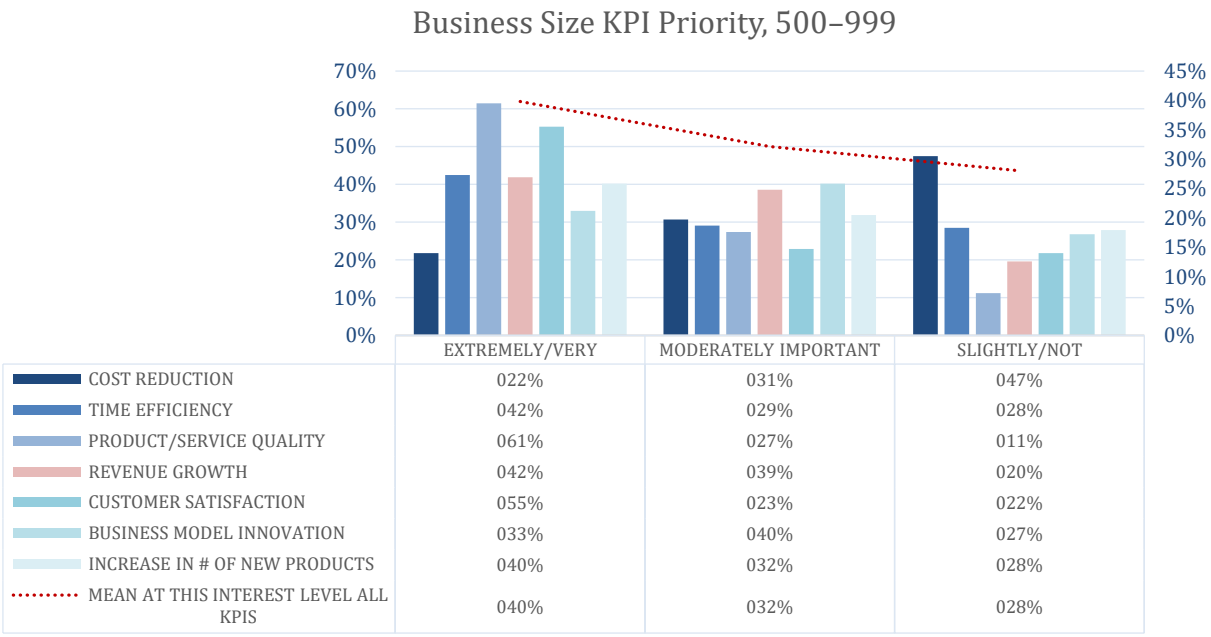
Time efficiency is significantly less regarded as "extremely or very important" than in other sectors.

Companies with 250–499 Employees Segment Analysis

The small/midsize companies that have invested in or are considering investing in BDTs are focused on growth. BDA can enable them to analyse their existing customer base and pricing to identify upselling opportunities and, critically, to provide market analysis to identify new customers.

Diversification is also a key growth opportunity, and analysis of the existing and potential customer base can support the critical decision to invest in developing new or modified products and services.

3.3. Importance of KPIs in Companies with 500–999 Employees



	EXTREMELY/VERY	MODERATELY IMPORTANT	SLIGHTLY/NOT
COST REDUCTION	022%	031%	047%
TIME EFFICIENCY	042%	029%	028%
PRODUCT/SERVICE QUALITY	061%	027%	011%
REVENUE GROWTH	042%	039%	020%
CUSTOMER SATISFACTION	055%	023%	022%
BUSINESS MODEL INNOVATION	033%	040%	027%
INCREASE IN # OF NEW PRODUCTS	040%	032%	028%
MEAN AT THIS INTEREST LEVEL ALL KPIS	040%	032%	028%

Figure 18 — Importance of KPIs in Large Enterprises (500–999 Employees)

Source: DataBench Survey, October 2018, 700 interviews, sample size n = 179

Companies with 500–999 Employees Higher Priorities

An increase in the number of new products or services launched is significantly more regarded as "extremely or very important" than for other organisation size segments.

Revenue growth is significantly less regarded as "slightly important or not important" as a goal for BDA adoption than in other organisation size segments.

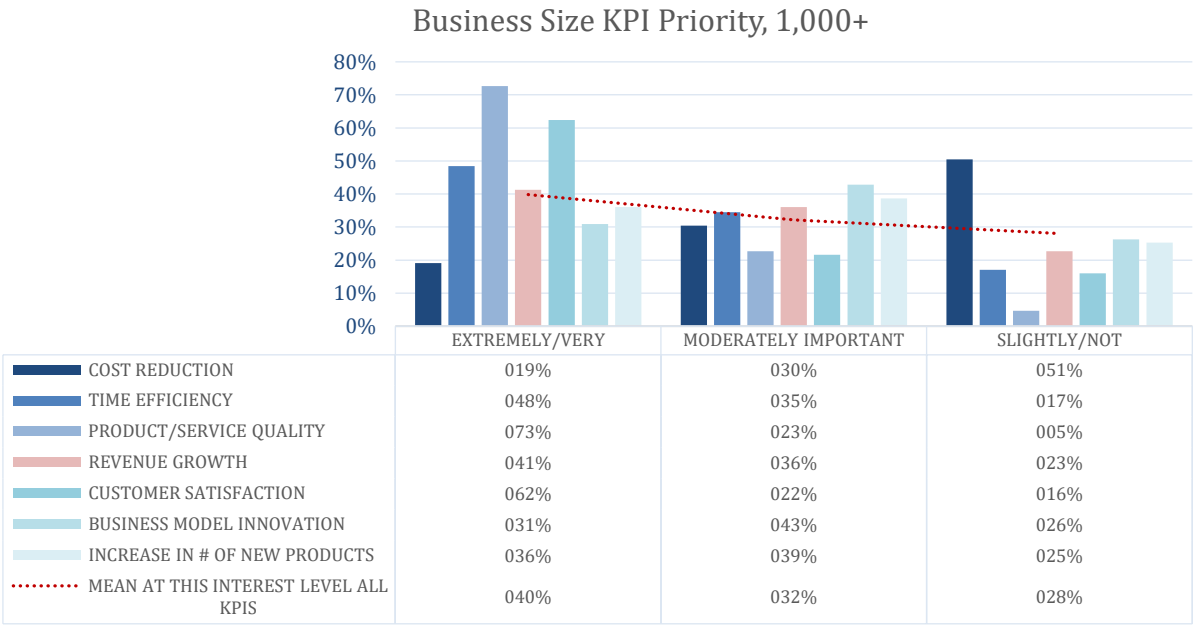
Companies with 250–499 Employees Lower Priorities

There are no KPIs in this sector that are significantly less important than in other sectors.

Companies with 250–499 Employees Segment Analysis

Medium-sized companies are often in a better position than smaller/start-up companies to consider investment in new products and services, whether via internal R&D, mergers and acquisitions, or OEM and reseller partnerships. Effective BDA of existing internal product sales and the financial and market prospects of potential M&A targets or partners is a valuable decision support tool. Detailed analysis of market prospects is likely to be expected from external partners or investors.

3.4. Importance of KPIs in Companies with 1,000+ Employees



	EXTREMELY/VERY	MODERATELY IMPORTANT	SLIGHTLY/NOT
COST REDUCTION	019%	030%	051%
TIME EFFICIENCY	048%	035%	017%
PRODUCT/SERVICE QUALITY	073%	023%	005%
REVENUE GROWTH	041%	036%	023%
CUSTOMER SATISFACTION	062%	022%	016%
BUSINESS MODEL INNOVATION	031%	043%	026%
INCREASE IN # OF NEW PRODUCTS	036%	039%	025%
MEAN AT THIS INTEREST LEVEL ALL KPIS	040%	032%	028%

Figure 19 — Importance of KPIs in Very Large Enterprises (1,000+ Employees) (% of respondents)

Source: DataBench Survey, October 2018, 700 interviews, sample size n = 194

Companies with 1,000+ Employees Higher Priorities

Time efficiency is significantly more regarded as "extremely or very important" than in other organisation size segments and is significantly less regarded as "slightly important or not important".

Product/service quality is also significantly more regarded as "extremely or very important" than in other organisation size segments and is significantly less regarded as "slightly important or not important".

Customer satisfaction is also significantly more regarded as "extremely or very important" than in other organisation size segments and is significantly less regarded as "slightly important or not important".

Companies with 1,000+ Employees Lower Priorities

Cost reduction is significantly more regarded as "slightly important or not important" as a goal for BDA adoption than in other organisation size segments and less regarded as "extremely or very important".

Companies with 1,000+ Employees Segment Analysis

Large companies have the resources to invest in BDA, and BDTs are now able to integrate information that was previously in silos to enable consolidated views of both the company's own processes, which can be optimised, and the current or potential market for products and services.

Large companies have access to far more internal customer data, and the increasing ability to integrate this leads to significant opportunities in improving customer relations, while at

the same time upselling and cross-selling to those customers with carefully targeted offerings.

3.5. Concluding Remarks

This chapter has analysed the survey results on the relative importance by company size of business KPIs used for benchmarking of BDTs, as presented in the previous deliverable, D.2.2 (from which the Figure 20below is sourced), to investigate users' main needs as expressed in terms of their priorities.

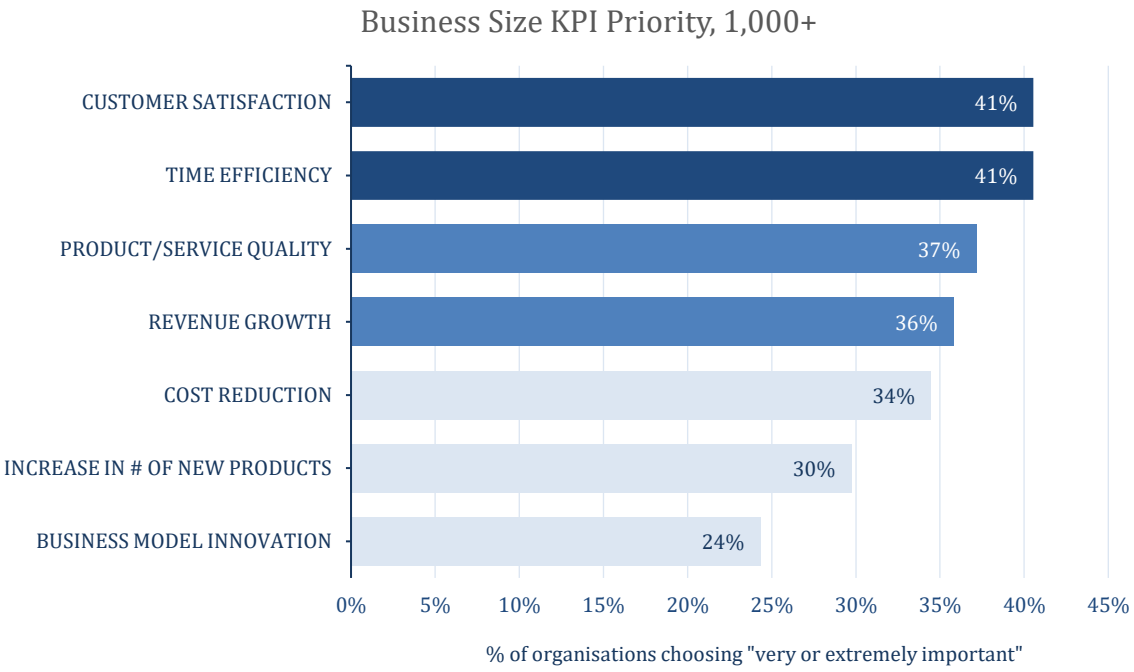


Figure 20 — KPI Ranking by Importance for SMEs

Source: DataBench Survey, October 2018, 700 interviews

As noted in section 3.1, further comparison with the other business size categories indicates that the significantly different KPI priorities for SMEs are that cost reduction is considered much more important as a BDA KPI, while product/service quality, revenue growth and customer satisfaction are significantly less important. The focus on the cost reduction KPI is worth noting because this KPI is considered the least significant across all sectors (with minor exceptions in manufacturing and transport/logistics).

This difference in emphasis can be better illustrated visually by comparing SME and non-SME respondents' views on the importance of the KPIs in benchmarking the impact of BDA adoption.

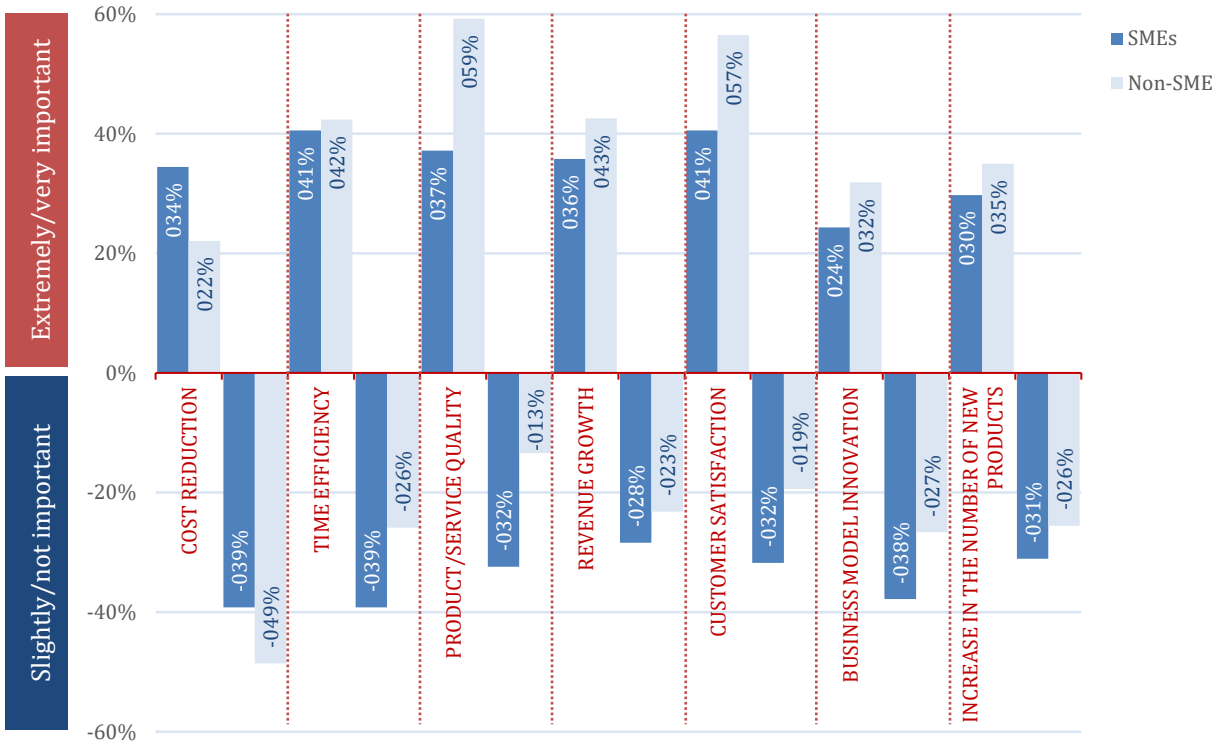


Figure 21 — Comparison of Importance of KPIs in Benchmarking BDA Adoption Between SMEs and Non-SMEs

Source: DataBench Survey, October 2018, 700 interviews (positive bars are responses for "extremely/very important"; negative bars are for "slightly/not important")

Although it is entirely understandable that cost control is vital in the SME sector, whether for start-up companies or mature SMEs without significant growth aspirations, it is possible that encouraging BDA implementations with more transformative potential could have a more significant effect on the very important SME economic sector.

For example, large chain food retailers spend heavily on BDTs to maximise customer satisfaction, and so market share. With the growing availability of sector-wide pricing and consumption data, smaller food retailers can also make greater use of pricing and consumer preference trends analysis.

As noted, there seems to be a parallel opportunity in the agricultural sector, where there is also a significant gap between BDA utilisation between large companies and the more typical smaller producers.

In general, if BDA product and service providers were able to provide more affordable packaged and targeted BDA solutions to SMEs, proposed as insight as a service, SMEs would have the opportunity to apply BDA more effectively, while also creating an expanded market for BDA providers.

For small to midsize companies, where revenue growth is the prevailing priority, BDA products and services specifically designed to identify opportunities for market diversification, and tailored to sector priorities off the shelf, will appeal directly to the BDA adoption priorities of successful and growing organisations.

As such companies do grow successfully, as illustrated by the responses from the 250–499 employee segment, the number of new products or services launched becomes the top

priority to widen existing market offerings and/or enter new markets. BDA offerings to inform internal R&D or expansion via partnership or M&A will address the primary needs of these economically important engines of growth.

4. The Survey of ICT Projects

DataBench objectives include the analysis of the H2020 ICT 14 and 15 projects and specifically the interaction with industrial partners running BDT trials to investigate their business KPIs and to share the results from the projects' business benchmarking. This is part of the DataBench activities to interact with the research and benchmarking community and to pave the way towards the production of a Benchmarking Handbook aligned with the users' needs.

To do so, we will launch a second wave of the DataBench survey inviting the industrial partners of these projects to take the survey and possibly agree to share their results by participating in case studies (implemented under WP4). The second wave of the survey will take place from June 2019.

In this context, this chapter provides an overview of the Horizon 2020 projects we plan to reach out to and their trials and pilots, classifying them in the main industries examined in this report. The idea is to help the industrial partners of these trials to identify the survey results relevant to them, particularly concerning the business KPIs.

The chapter provides also a description of the self-assessment tool that we have developed in integration with the second wave of the survey. This is a real-time personalised report that all respondents will receive on completion of the survey, providing a comparison of their survey answers and those of their peers (the other respondents belonging to the same industry and company size class). The goal is to incentivise participation in the survey and to provide immediate feedback to respondents, which can be used as a starting point of an eventual case study.

4.1. The H2020 Projects Targeted by DataBench

To further contextualise and validate the results provided in D2.2 and in this deliverable, DataBench aims to issue a second wave of the survey to the general public and to H2020 projects engaged in Big Data projects (ICT 14 and 15). Other Horizon 2020 projects will also be considered for the survey. As seen in the table below, a number of projects will be addressed in the new survey. In addition to the survey, WP2 will provide every respondent with a real-time personalised report benchmarking the business choices and KPIs against their peers.

ICT #	Project Name	ICT #	Project Name
ICT-14	Big Data Ocean	ICT-12	CloudButton
	AEGIS		ELASTIC
	euBusinessGraph		EXA MODE
	QROWD		ExtremeEarth
	FashionBrain		INFORE
	EW-Shopp		SmartDataLake
	DataPitch	ICT-13	MUSKETEER
	BodyPass		Safe-DEED
	SLIPO	ICT-16	BigDataStak
	EDI		E2DATA
	Lynx		Track and Know
	TheyBuyForYou		BigDataGrapes
	FANDANGO		CLASS
	Icarus		I-BiDaaS
	Cross-CPP		Typhon
ICT-15	DataBio	ICT-18	SODA
	TT: Transforming Transport		MH-MD
	BigMedilytics		SPECIAL
	BOOST 4.0		

Table 3 — Horizon 2020 Projects Targeted with the Second Wave of the Survey

Source: IDC elaboration of H2020 data

The following section provides a classification of the ICT projects targeted in the vertical market survey. However, the analysis of use cases and business KPIs (Chapter 2) is performed only for those projects that concentrate on developing technologies and solutions for one single vertical market. Hence, the "vertical ICT projects" definition.

The analysis offers an in-depth understanding of the Horizon 2020 projects (and pilots and use/business cases). The analysis will correlate each vertical project to specific industry KPIs and use cases. In doing so, it will be possible to estimate and evaluate the objectives, the needs and the industry relevance of each project. In the next steps of the project, we plan to integrate into the database the responses given by the respondents of the second wave of the survey.

4.1.1 Methodology to Classify Projects

For an aggregate and preliminary analysis of the most common pilots and use/business cases by vertical, IDC uses a NACE-based classification system. To comply with the quota requirements, the verticals have been grouped into nine macro sectors (table below). Given the purpose of the study, the classification does not consider three other verticals — consumer, government, and personal and consumer services.

Resource Industries	Financial Services	Manufacturing	Retail and Wholesale	Transport & Logistics
Agriculture	Insurance Banking Securities and investment services	Discrete manufacturing Process manufacturing	Retail Wholesale	Transport
Professional Services	Telecom/ Media	Utilities/Oil and Gas		
Business/ IT services	Telecom Media	Utilities	Healthcare	

Table 4 — List of Industries in DataBench Survey

Source: IDC, 2018

To classify the projects and use/business cases, IDC's vertical experts adopted the "end-user" classification according to the website projects and pilot and use/business case description. With the end-user classification, the solutions provided in each project or pilot are related to the specific final user.

To classify the projects, we have used an end-user view, resulting in the sum of the whole verticals addressed in the project pilots.

4.1.2 Analysis

As discussed, the analysis only covers the projects that feature an exclusive vertical solution/technology. This is to focus KPIs and use case analysis on specific and targeted solutions that aim to use BDA to improve niche projects in specific vertical sectors.

	Project	IDC Verticals								
		Resource Industries	**Financial Services**	**Manufacturing**	**Retail and Wholesale**	**Transport & Logistics**	**Professional Services**	**Telecom/ Media**	**Utilities/Oil and Gas**	**Healthcare**
ICT-14	Big Data Ocean					1	1		1	
	AEGIS		1							
	euBusinessGraph	1	1	1	1	1	1	1	1	1
	QROWD					1				
	FashionBrain				1					
	EW-Shopp	1	1	1	1	1	1	1	1	1
	SLIPO	1	1	1	1	1	1	1	1	1
	Lynx	1	1	1	1	1	1	1	1	1
	TheyBuyForYou	1	1	1	1	1	1	1	1	1
	FANDANGO							1		
	Icarus	1		1		1			1	
ICT-15	DataBio	1								
	TT: Transforming Transport				1	1				
	BigMedilytics									1
	BOOST 4.0			1						
ICT-12	CloudButton					1	1			
	ExtremeEarth	1					1			
ICT-13	MUSKETEER			1						1
ICT-16	BigDataStak		1		1	1				
	E2DATA		1		1			1	1	1
	Track and Know		1			1				1
	BigDataGrapes	1		1						
	CLASS					1				
	I-BiDaaS		1	1				1		
ICT-18	SPECIAL		1			1				

	Horizontal ICT project — the scope of the pilots/UCs within this ICT project is not sector specific, so we define this as a technology-specific project (e.g., the technology is applicable to all the verticals)
	Cross-sector ICT project — the scope of the pilots/UCs within this ICT project covers a group of sectors (specifically, different pilots in the project target different industries)
	Vertical ICT project — the scope of the pilots/UCs within this ICT project covers only one vertical (the solution developed by this ICT project is industry specific and all pilots address the same industry)

Table 5 — Classification of Horizon 2020 Projects by Verticals Targeted

Source: IDC elaboration of H2020 data

Agriculture: DataBio

DataBio (Data-Driven Bioeconomy) aims to deliver a Big Data platform to ease and strengthen the cooperation among end users and vendors, technology institutes and other organisations from the bio economy sector. This will be achieved by developing a solution that demonstrates how to produce food, energy and biomaterials sustainably with "agricultural" Big Data. The pilots focus on "Agriculture" (to build a geo-coded map of agricultural fields and provide real-time monitoring of activities to increase efficiency of resource use), "Fishery" (to control small pelagic fisheries in the North Atlantic Ocean and the tropical tuna fisheries) and "Forest" (to state forest conditions — health and damages

— and create a recommendation tool for planting management plans). These fall into the four industry-specific use cases: precision agriculture, yield monitoring, field mapping and crop scouting, and heavy equipment utilisation.

As seen in Chapter 2, the potential of BDA and BDTs in the agricultural sector has yet to be fully explored. European agricultural companies are perhaps slightly behind their worldwide peers for several reasons. Firstly, European companies are smaller in terms of employees and availability of fields, especially when compared against their peers in North America. Secondly, European organisations are traditionally family owned and managed. The capital available for BDA is limited and for long time Big Data hasn't been perceived as a game changer to improve agriculture productivity. Big Data (and data in general) has been mostly used in R&D to improve quality and resistance of seeds and plants, including gene research.

With this project, the consortium is trying to change the traditional ways. The use of bio economy data to map agricultural fields and fisheries delivers insights to improve productivity of fields and optimise fishing activity, focusing only on those shoals of fish that are ready to be caught to protect the sustainability of the species. Considering the third pilot on forest data platforms, the consortium engages in the well-being of forests and green areas and implementation plans.

Financial Services: AEGIS

AEGIS (Advanced Big Data Value Chain for Public Safety and Personal Security) is focused on creating an interlinked public safety and personal security data value chain. It aims to deliver a platform through which users can address some of the more burdensome data processes, such as data curation and integration, analysis and intelligence sharing. Within this project the consortium is working to strengthen a data-driven mentality in European companies. Alongside these benefits, organisations will be able to increase the value of their data sets, leveraging the opportunities created by the data sharing economy. Of the three pilots developed in the project, only one ("insurance sector: support, warning and personal offering") is relevant for our analysis due to the decision to focus only on business-related solutions, and not domestic and government-related solutions. The pilot aims to offer and support personalised solutions for asset management, forecasting and near-real-time event detection and analysis, with the aim of improving both customer satisfaction and cross-selling/up-selling opportunities for insurance companies.

As seen in Chapter 2, the project is in line with the expectations for Big Data use in the financial service vertical. The main KPI for financial and insurance organisations is to improve customer satisfaction, broadening both offerings of new products and services but also leveraging the opportunities to cross-sell and up-sell opportunities. More informed insight into customer needs and desires, when properly leveraged, can reduce the cost to deliver personalised services. In the insurance sector, forecasting and prediction solutions can create real-time tailored insurance policies based on customer data to forecast customer risk and offer the right policy price. The objective of the project matches the specific usage-based insurance case studies identified in this study, as well as more common use cases such as price optimisation, new product development and customer profiling, targeting and optimisation of offers.

Manufacturing: BOOST 4.0

BOOST 4.0 (Big Data Value Spaces for Competitiveness of European Connected Smart Factories 4.0) aims to create an efficient and reliable European industrial data space to strengthen the competitiveness and opportunities of Industry 4.0, leveraging the Big Data (and analytics) within European factories. The project aims to:

- Create global standards and contribute to the international standardisation of European industrial data space data models and open interfaces
- Secure digital infrastructures through the extension of cloud and edge digital infrastructures to ensure high-performance operations
- Offer trusted Big Data middleware integrating the four main European OS initiatives
- Develop digital manufacturing platforms to develop data pipelines for advanced analysis and data visualisation in open source environments
- Deliver certification programmes for equipment, infrastructure, platforms and Big Data services

The project involves 10 Europe-based factories as case studies to see how Industry 4.0 can be implemented in Europe. The factories belong to Volvo (Finland), Fill Your Future (Austria), Philips (the Netherlands), Bentler (Germany), Riastone (Portugal), GF (Switzerland), Whirlpool (Italy), CRF (Italy), Gestamp (Italy) and Volkswagen (Germany). The project aims to enhance the ability to integrate the increasing amount of data across heterogeneous data sources, enabling European organisations to create a unique data space to monitor and analyse data across all lines of business, from procurement to final delivery to customers.

BOOST 4.0 also covers use cases identified for the manufacturing sector, such as supply chain management, smart warehousing and predictive maintenance. Leveraging Big Data for an in-depth study of the three use cases will enable the project to create an efficient data-driven connected smart factory to improve European manufacturers' competitiveness.

Retail and Wholesale: FashionBrain

FashionBrain (Understanding Europe's Fashion Data Universe) is developing a tool to combine data for and from the fashion industry from heterogeneous data sources. With this technology it will be possible to forecast upcoming fashion trends (social media analysis) and provide effective and personalised recommendations for customers and retailers. For instance, retailers will be able to optimise offerings and procurement processes according to what is requested by the market and customers. The ICT project has two main lines of development. From a customer point of view, the tool will help users to identify a specific garment or accessory from a picture and directly match it to retailers' product catalogues (Shop the Look). From a business perspective, the tool will help retailers to optimise their supply chain using accurate predictions of future fashion trends (Fashion Trend Prediction), for example, analysis of influencers' profiles. The project tackles four of the industry use cases identified by the DataBench survey: customer profiling and targeting and optimisation of offerings, automated customer service, supply chain optimisation, and inventory and service (parts) optimisation. In developing solutions in these use cases, the project scope ties in perfectly with the DataBench objectives.

Relating the project to the business analysis in Chapter 2, FashionBrain is trying to improve the business KPIs identified as essential by the survey respondents. BDA can help to optimise retailers' supply chains, predicting market trends and consumers' preferences, and reducing the risk of product stock-out. Once the tool is ready, it will also be a solution for business model innovation. With the technology, there will be a paradigm shift in how retailers can be reached by consumers, improving the visibility of products also for smaller retailers.

Transport and Logistics: QROWD and CLASS

QROWD (Because Big Data Integration is Humanly Possible) aims to improve and better leveraging cross-sectoral streaming of Big Data for integration, across heterogeneous data sources and data types (e.g., geographic, transport, meteorological and news data) capitalising on human feedback. The project considers real-time integration of data for urban mobility and infrastructure information to engage in urban data-driven innovation and prediction of road conditions. For our purposes, the only pilot suitable for analysis (the other two are focused on the government vertical) is "Traffic Service", which aims to implement a generalised traffic report solution to understand the weaknesses and strengths of cities' mobility. The pilot looks at how to better track and regulate changes in traffic conditions. From a use case perspective, the project doesn't specifically fit into any of the identified use cases because it is focused on traffic management (while the identified use cases are tailored for the logistics vertical).

It is also not worth matching the project with the business KPI analysis as the project focuses on improving customer usage and satisfaction in mobility. Indeed, this project develops services classified in the transport/logistics vertical, but addressing the smart city potential market, which IDC considers as mainly a government market.

Within the transport/logistics vertical, CLASS (Edge and Cloud Computing: A Highly Distributed Software Architecture for Big Data Analytics) is an ICT 16 project that aims to create a technical solution to effectively couple data in motion and data at rest, creating an architecture that can work within the compute continuum (from edge to cloud) and provide real-time insights. The pilot projects are "Intelligent Traffic Management" and "Advanced Driving Assistance System — ADAS". The former deals with traffic lights and smart road signals ("green routes" for emergency vehicles and traffic enhancement through intelligent road management), while the latter is based on intelligent road management and obstacle detection and general city information.

Telecom/Media: FANDANGO

FANDANGO (Fake News Discovery and Propagation from Big Data Analysis and Artificial Intelligence Operations) aims to aggregate different data sources and information (social media, open data, news, etc.) at a European level to provide European citizens with more reliable information — eliminating "fake news". The objective will be achieved by lowering data interoperability barriers with the creation of a unified Big Data platform to support both traditional and modern media industries in the sourcing of information. The project aims to reduce fake news in three areas: climate (climate change, natural disasters, etc.), immigration (reducing misleading information on immigration) and European context (reducing unreliable news that might undermine European integration and democratic processes). This is a niche project for the EU, so there is no direct match with the specific use cases identified for the telecom and media sector (network analysis, ad targeting and

scheduling optimisation, etc.) or with more common use cases (customer profiling, targeting and offering optimisation, etc.).

Similarly, it isn't possible to match the scope of the projects with the most important business KPIs for this sector. Among the main business KPIs identified, only "improvement of product/service quality" can be associated with the project. The main objective of the three pilots is to create a more informed and conscious (higher quality) information service for European citizens. However, the project's scope is far removed from business considerations as it focuses more on the social aspects of the telecom/media industry, so pure business KPIs are not relevant for the analysis.

Healthcare: BigMedilytics

BigMedilytics (Big Data for Medical Analytics) aims to reshape the European healthcare system through the deployment of state-of-the-art BDTs to strengthen the connections and links between heterogeneous patient data sets to improve customisation of treatments and solutions. In terms of objectives, the project identifies the delivery of high-quality care and cost reduction. It also aims to strengthen the collaborative innovation environment across European public and private healthcare providers. The project identifies three main lines of intervention (pilots) to be studied. The first is "Population Health and Chronic Disease Management" and aims to reduce the cost of the healthcare system through greater personalisation of treatment. The second is "Oncology" and this focuses on reducing the economic burden of cancer treatments and the complications that come with the most common cancers (prostate, lung and breast). The third pilot is the "Industrialisation of Healthcare Services", which aims to optimise processes and management of more traditional and common workflows such as stroke management, sepsis management and radiology workflows. The three pilots match all the industry-specific use cases reported in the DataBench survey. The use cases under consideration are illness/disease diagnosis and progression and personalised treatment via comprehensive evaluation of health records and quality of care optimisation. The BigMedilytics solution is aimed at infusing greater value into the targeted processes and operations through the deployment of BDA and BDTs.

From the analysis of the business needs and KPIs, the project's priorities are in line with the average evaluation of business priorities. The analysis highlights a strong interest in customer satisfaction, which in medical terms means the resolution of the illness or the reduction of re-admissions for the same disease. Customer satisfaction can also be achieved through improvements in the customisation of treatments and medicaments and a reduction in the recovery rate.

4.2. The Self-Assessment Tool

Within the WP2 framework, IDC has created the self-assessment tool, a benchmarking solution on the usage of BDA and BDTs and the benefits of BDA on the most important business KPIs.

4.2.1. What it Is

The self-assessment tool is an interactive web-based tool implemented as an add-on to the DataBench survey and sent to respondents in a report format (PDF file), comparing the respondents' answers with the current data set. The tool is a simple visualisation solution to benchmark seven questions from the DataBench survey. ICT projects, those taking part in the WP4 interviews and the general public will benefit from the final report, as it shows

their BDA and BDT position against their European peers. In the customised report, "personal" responses to the Big Data survey are compared with responses from peers in the same respondent's industry and company size to help to validate or reframe the BDT project needs. The report includes charts and written analytic feedback on current BDTs.

4.2.2. How it Is Built

The self-assessment tool provides initial guidance for companies to help them understand their current development and positioning of the BDA and BDT environments, and the benefits. The tool is built on IDC's DataBench survey, carried out in September–October 2018. As already highlighted, the survey aims to understand expectations and experiences in Big Data projects to better assess users' focus and needs. The survey is based on a sample of European business organisations in 11 Member States, resulting in 700 interviews segmented as follows:

- 11 Member States: France, Germany, the Netherlands, the UK, the Nordics (Denmark, Sweden), South Europe (Italy, Spain), CEE (Czech Republic, Poland, Romania)
- 16 industry sectors and 7 employment size classes
- Once respondents complete the survey, an automated data elaboration process is started to create a PDF report comparing the participants' answers against the database of the 700 responses to the initial DataBench survey. The report is provided in real time and can be downloaded for later consultation.

The report presents answers to the seven questions in the survey, with a short commentary of the respondents' positions against their peers. The peer category is identified by looking at the responses of the participants to the DataBench survey within the same industry sector and the same business size.

The questions analysed are business orientated and specify the reasons for the adoption of BDA, the importance and level of achievement of business KPIs, and interest in adopting BDA in a long-list of business use cases. The report is based on the following questions:

1. Which of the following business goals are driving the adoption or consideration of Big Data and analytics in your organisation?
    - Possible answers: customer service and support; engineering; research and development; product innovation (new business initiatives); maintenance and logistics; marketing; finance; HR and legal; sales; product management; governance, risk and compliance; IT and data operations; others; all the above.
2. How important are the following business KPIs for measuring the impact of your organisation's Big Data and analytics efforts?
    - The answer is given on a scale of "not important at all" to "extremely important" for the following KPIs: cost reduction, time efficiency, product/service quality, revenue growth, customer satisfaction, business model innovation, an increase in the number of new products or services launched
3. What level of benefits has your organisation achieved so far, or expects to achieve, from the use of a Big Data and analytics environment?
    - The answer is given on a scale of "negative impact" to "high level of benefits".

4. In percentage terms, what is the actual benefit realised, or which benefit do you expect to realise, from the use of Big Data and analytics for the following business KPIs?
    - The answer covers three business KPIs: increased profit, increased revenue and reduced cost. Respondents are given these pre-defined metrics: less than 5%, 5%–0%, 10%–24%, don't know.
5. To what extent has your organisation's deployment of Big Data and analytics been impacted, or will your organisation's deployment of BDA be impacted, by the ability to attain the following KPIs?
    - Respondents should say if they have seen a decrease, no change, slight increase, moderate increase or strong increase in the following KPIs: time efficiency, product/service quality, customer satisfaction, business model innovation, number of new products or services launched.
6. For the following business KPIs please estimate what percentage of expected improvement will be linked to the adoption of Big Data and analytics in 2020.
    - Respondents should quantify the expected improvement of a short-list of business KPIs: cost reduction, time efficiency, product/service quality, revenue growth, customer satisfaction, business model innovation, number of new products/services launched. They are given pre-defined metrics between 0% and 100%.
7. What is your organisation's position on each of the following specific Big Data and analytics business use cases?
    - Respondents should say if they are using or implementing, evaluating or planning, are interested but have no plans, are not interested and have no plans for the list of use cases used in the survey (see Annex for details).

4.2.3. How it Works

The tool is very user-friendly. The respondent accesses the survey through a link provided by a DataBench partner or directly from the DataBench website (for the general public). It then takes 20–25 minutes to answer the 20 questions in the survey. Once the questionnaire is completed, an email is sent to the respondent with a link to the downloadable version of the self-assessment report (survey landing page and email in the Annex).

From a developer's point of view, the creation of the self-assessment tool is slightly more complex.

The first step is the creation of a custom ICT project link to be sent to the targeted respondent (the ICT coordinator), with a request to share the link within the organisation and among the partners in the ICT project targeted. Once the receiver completes the survey, the self-assessment tool evaluates the responses and generates the benchmarking for each of the questions mentioned in section 3.1.2. The tool creates two sets of answers, one that assesses and benchmarks the respondent against peers in the same industry sector, and one that performs the same task but benchmarks the respondent against peers within the same business size.

At the same time, the responses are anonymised and will be used to update the database from which the tool gathers the information for the comparison with the peers. The frequency of the update will be defined at a later stage, as it will depend on the frequency and effective up-take of the survey by the respondents.

4.2.4. Objectives and Targeted Audience

The self-assessment tool is tailored to target business users specifically within:

i. The consortium of the ICT 14-15 projects. They will receive a custom link specifically named after the ICT project targeted by WP2. WP2 aims to target mostly the business partners within every project to get more precise and reliable answers on the business KPIs and performances. In addition to the ICT 14-15 consortiums, other ICT projects — mainly 12, 13, 16 and 18 — will be addressed to broaden the reach of the survey and to include further results.

ii. The general public. A link to the survey is also available on the DataBench website and is open to all interested organisations.

The self-assessment tool is designed to help achieve the specific objectives of WP2. The first objective is to promote and incentivise ICT project partners to complete the survey and to be DataBench case studies (WP4). WP2 members believe the self-assessment tool is an effective incentive for ICT project consortiums to engage in the survey and potentially become case studies for WP4. The second objective is to increase the number of responses, both among the ICT partners and the general public, to enrich and update the survey database and periodically update and validate the KPIs. The third objective is to use the data obtained to assess the case studies and demonstrate the range of the potential impacts/benefits delivered by the adoption of BDA solutions and technologies.

4.3. The Report

The self-assessment tool is an addition to the original survey created and carried out by IDC in the first six months of the project. This second version, as anticipated, will provide more detailed traceability of the respondents and the ICT project in which they are part of (confirmation of the ICT project and pilot and use/business case), alongside additional company information (name and surname of the respondent, email, company name). The improved traceability will enable deeper analysis of the ICT projects to better understand the BDT environment in Europe. The request for further contact details will enable the consortium to recruit the business partners of the ICT projects willing to participate in the DataBench case studies proposed and managed by Politecnico di Milano.

The report is a PDF document of about 12 pages, with charts showing the positioning of the respondents against their peers. Alongside the charts, there is a general and comparative description to better assess respondents' environment within the general landscape of their peers.

Figure 22 — Draft Cover Page of the Self-Assessment Report

The first page, as in Figure 22, provides details such as the DataBench logo and description and the logos of the consortium organisations. It also includes a summary of the respondent: the respondent and company name, the Horizon 2020 project in which they are involved and the survey completion date. This is followed by the benchmarking analysis, with the report covering the seven questions considered in section 3.1.2 for the benchmarking activity, analysing first the industry segment and then the size band.

4.4. Status and Next Steps

By the time this deliverable is published, the self-assessment tool will be operational and IDC will be in the process of launching the second phase of the survey, contacting the coordinators of the ICT projects targeted and asking for their support to reach out to their business partners. Here is the link to the survey on the Databench website page. Screenshots of the link are shown in the Annex: https://www.databench.eu/self-assessment-survey/

5. Technical and Business Choices

5.1. Overview

The Big Data technology (BDT) landscape is very complex due to the ever-increasing variety of technical solutions to Big Data issues, such as Big Data size (volume), different types of data to be matched and correlated (variety), and database storage and analytics process selection. Given these complexities, European organisations should find the right balance between BDA and BDT requirements, their actual needs, the solutions already in place and those available on the market. To do this, they should weigh up the business and technical needs to choose the right solutions to achieve the maximum level of business benefit and ROI.

This chapter outlines the analysis of the technical questions of the survey using two approaches. Section 5.2 provides a simple descriptive analysis of the highlighted questions, giving a general flavour of the BDT environment among European companies. The following section (5.3) provides an in-depth statistical analysis of the technical and business questions, to better understand correlations and common patterns in BDT adoption. Conclusions are provided in the closing section.

5.2. Descriptive Analysis of Technical Questions of the Industrial Needs Survey

This section presents a descriptive analysis of the technical questions from the industrial needs survey of 700 European businesses in 11 EU Member States.

5.2.1 Methodology

The companies participating in the survey confirmed their actual or planned use of Big Data. This analysis looks at how European companies are using BDTs and the technical challenges they face. Some questions investigate the current situation, while others focus on companies' planned initiatives. In particular, as detailed in Table 6, technical questions investigate the size of data in terms of data storage, the type of data currently stored and processed by the company, the type of data storage currently adopted, the current approach to data management, and the importance of each Big Data processing paradigm. A further set of questions investigates current and future (in two years' time) usage of technical performance metrics as well as the current and planned adoption of Big Data analytics.

Size of BD and Analytics	Type of Data	Type of Data Storage	Current Approach to Data Management
• Gigabytes • Terabytes • Petabytes • Exabytes	• Tables, files or structured data • Text data • Graphs or linked data • Geospatial or temporal data • Media (images, audio or video) • Time series (including IoT data) • Structured text (e.g., XML, genomic data)	• Relational database management systems (RDBMS) (e.g., Oracle 12c, IBM DB2, Microsoft SQL Server) • Columnar databases (e.g., SAP Sybase IQ, HP Vertica) • In-memory databases (e.g., SAP HANA) • NoSQL databases (e.g., MongoDB, Splunk) • Graph databases (e.g., Horton, Neo4j) • NewSQL databases (e.g., NuoDB, Google Spanner) • Hadoop • Open source Big Data platforms (not Hadoop) (e.g., Cassandra, Apache Spark) • Commercial Big Data platforms (e.g., 1010data, Extrahop) • Database appliances (preconfigured system with database management software, compute and storage resources, e.g., HP AppSystem for SAP HANA, Oracle Exadata)	• Structured and transactional data is captured and curated within a data warehouse or operational data store • Structured and unstructured data from different locations, including on-premise and in the cloud, is captured and curated across multiple repositories • Enterprisewide repositories or data lakes are used to capture, organise and curate data from multiple sources and formats across the organisation • Real-time data (e.g., log files, social media and IoT data) is streamed and used alongside other contextual data from your Big Data environment • The Big Data platform is available/exposed to external customers/partners and developers to build and extend data-driven applications

Data Processing Paradigm	Technical Performance Metrics	Analytic Techniques
• Batch processing • Stream processing • Interactive/near real time • Iterative/in-memory	• End-to-end execution time • Throughput • Cost (e.g., $/transaction) • Accuracy/quality/data quality/veracity • Availability	• Descriptive analytics • Diagnostic analytics • Predictive analytics • Prescriptive analytics

Table 6 — Summary of Technical Questions

Source: DataBench Survey, October 2018

5.2.2 Analysis of BDT Technologies in the Survey

First, the general results of this analysis have been taken into consideration to understand the general use of Big Data technologies in all European companies.

A more detailed industry analysis has also been conducted to verify whether there are more specific trends that characterise different industries. However, we could not find industry-specific evidence that significantly deviates from the insights obtained in the aggregate analysis. This report presents the overall summary results.

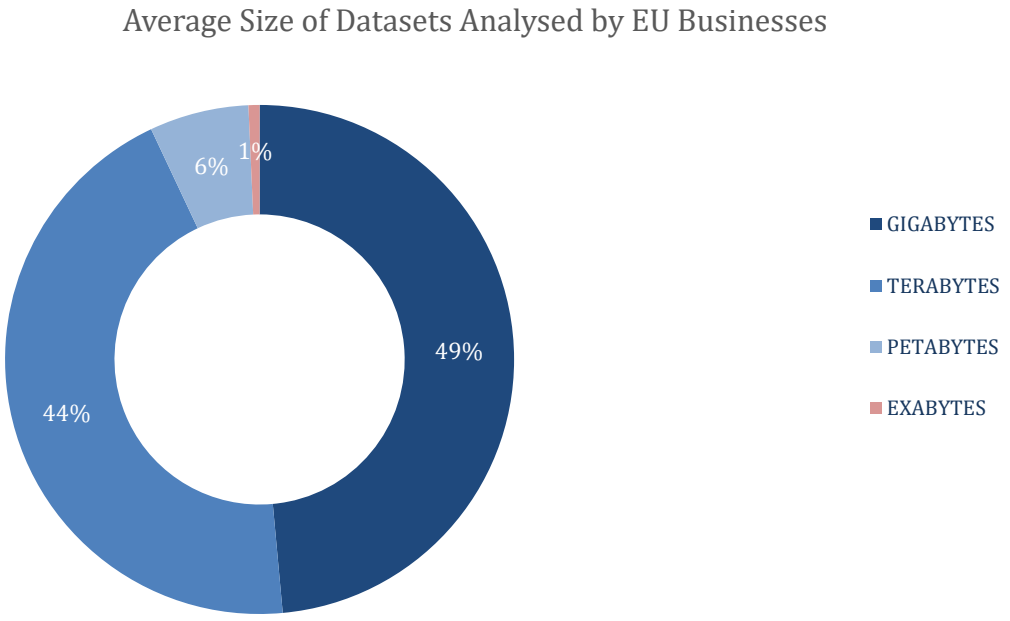


Figure 23 — Data Size (% of respondents)

Source: DataBench Survey, October 2018, 700 interviews

Figure 23 shows that European companies mainly analyse and store **gigabytes and terabytes of data**, while a small number of companies (less than 10%) deal with petabytes and exabytes. Nonetheless, data size is expected to grow significantly in the next few years, making companies' technical choices increasingly relevant.

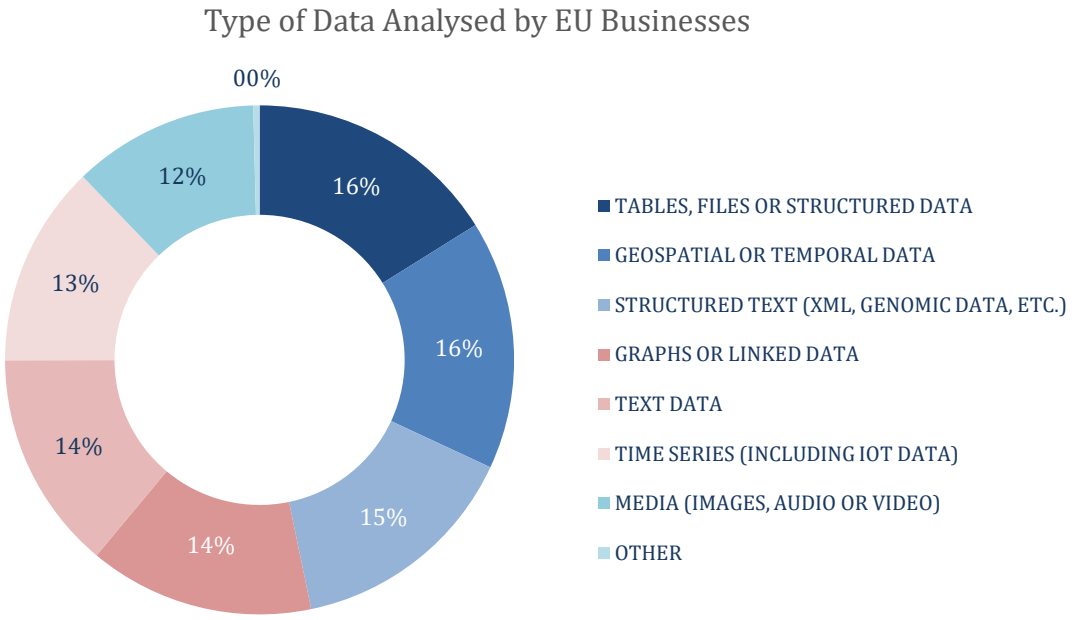


Figure 24 — Data Types (% of respondents)

Source: DataBench Survey, October 2018, 700 interviews

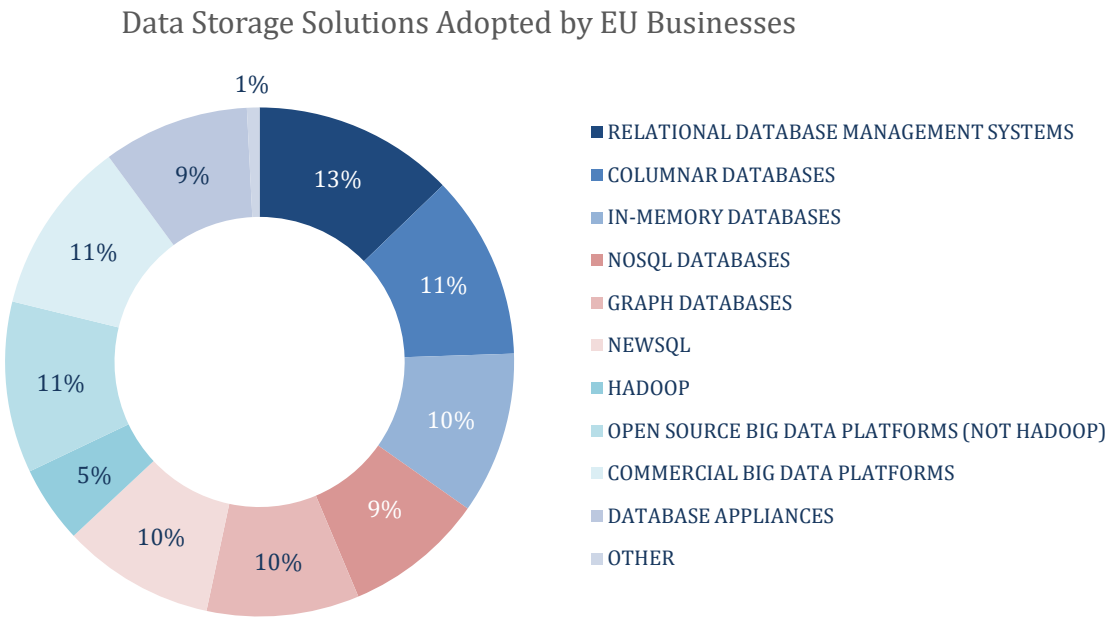


Figure 25 — Data Storage Solutions (% of respondents)

Source: DataBench Survey, October 2018, 700 interviews

As shown in Figure 24, EU companies process and store heterogeneous data. Tables and structured data account for a significant share of this data, followed by structured text and graphical data. Geospatial and temporal data, time series and media data are also important, however. These results highlight the emerging need to integrate heterogeneous data to effectively exploit all the information gathered by companies. The need to deal with this data is further emphasised by the adoption of a wide set of storage solutions (Figure 25).

However, RDBMS are the most widely adopted solution, given the importance of tables and structured data. From a storage perspective, relational solutions are followed by columnar and in-memory data stores; considering the selection of processing solutions, respondents indicate Hadoop and open source Big Data platforms as the most common environments. This is consistent with the approach to data management reported in **Error! Reference s ource not found.**. Indeed, around 50% of companies currently rely on data lakes to capture and organise data collected from multiple sources and formats across the organisation. Significantly, 16% of companies are pioneering the management and exploitation of real-time data.

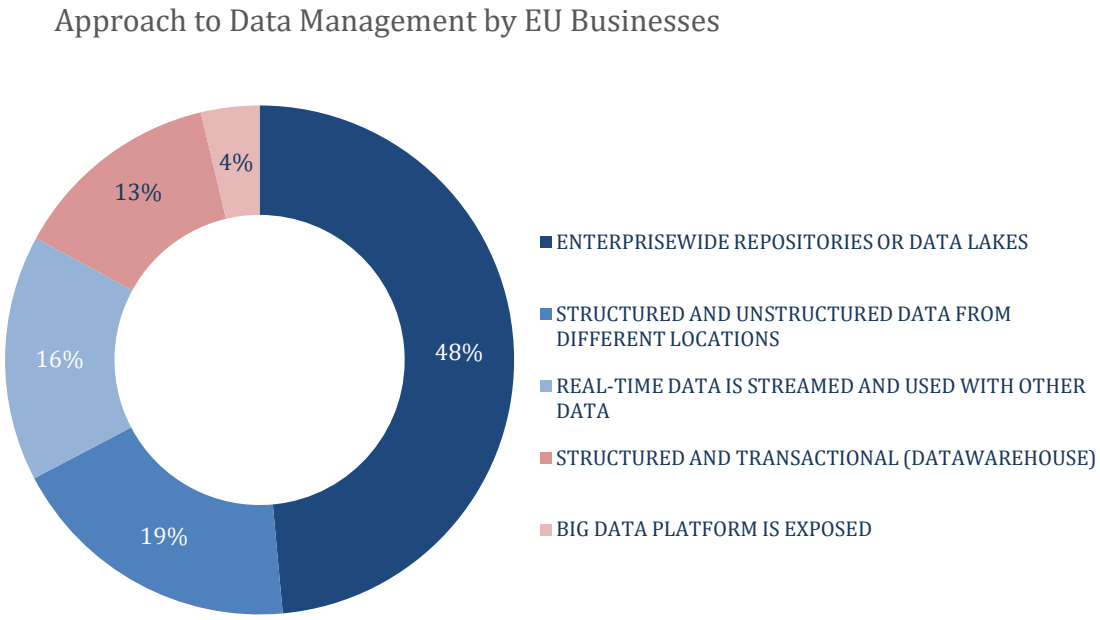


Figure 26 — Approach to Data Management (% of respondents)

Source: DataBench Survey, October 2018, 700 interviews

Figure 27 shows the data processing paradigms currently adopted by EU companies. The type of data processing has shown similar results for the different paradigms proposed in the survey. However, streaming and near-real-time data processing appear to represent emerging needs. Currently, the types of analytics most popular among European companies are descriptive and diagnostic analytics (Figure 28). In the future, companies are planning to move to prescriptive and predictive analytics. Currently, as shown in Figure 29, the most common technical performance metric is availability, along with accuracy, veracity and data quality, while end-to-end execution time, cost and throughput are less relevant. The last three metrics are expected to catch up in the next two years. It is worth noting that roughly 20% of interviewees are not using or planning to use any performance metrics.

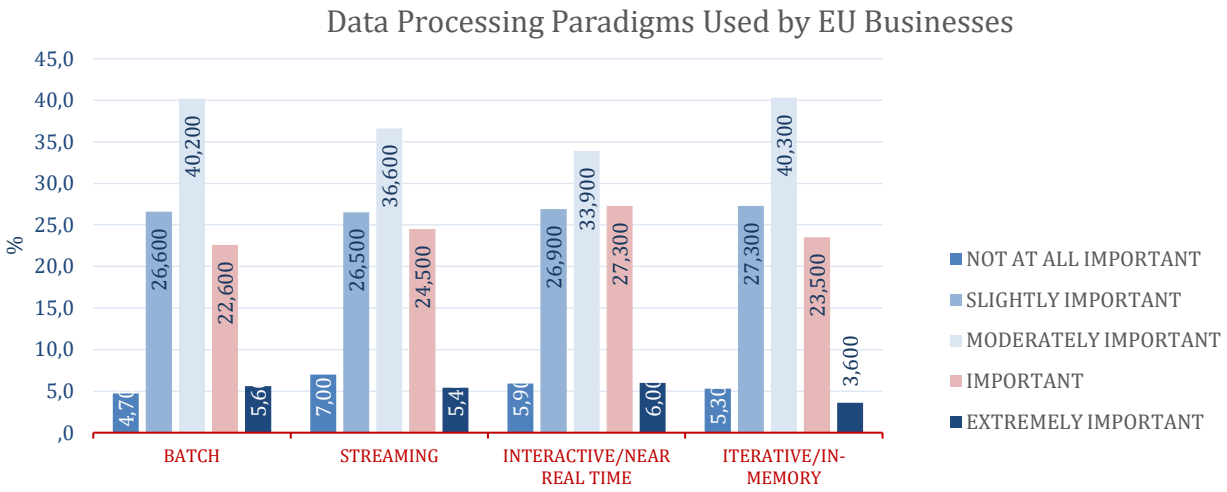


Figure 27 — Data Processing Paradigms (% of respondents)

Source: DataBench Survey, October 2018, 700 interviews (% of respondents, multiple answers)

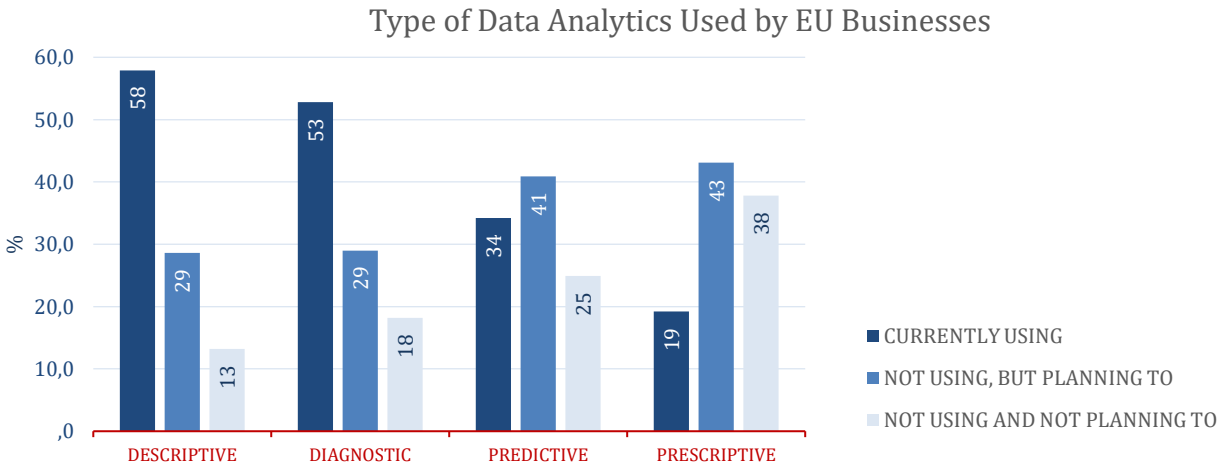


Figure 28 —Type of Data Analytics (% of respondents)

Source: DataBench Survey, October 2018, 700 interviews (% of respondents, multiple answers)

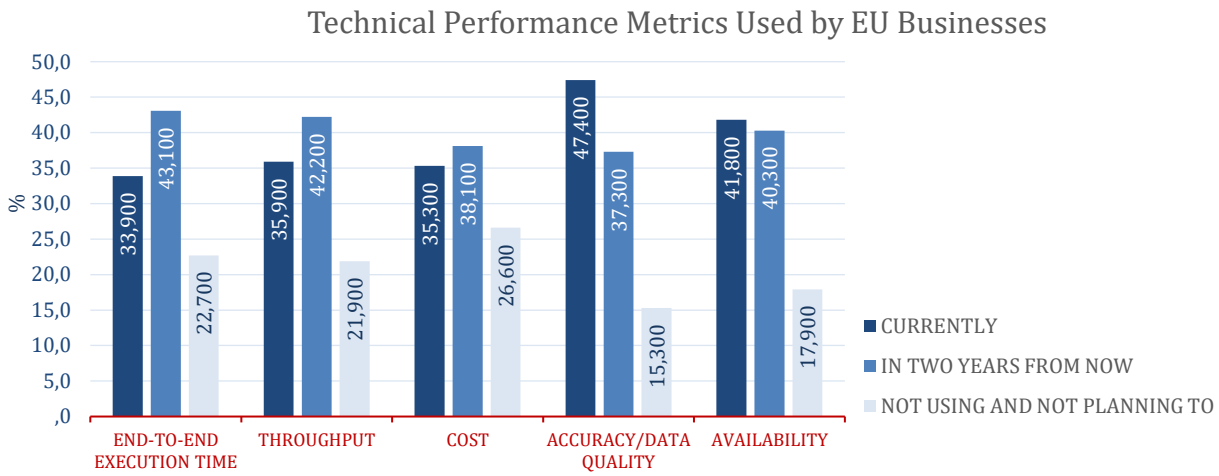


Figure 29 — Technical Performance Metrics (% of respondents)

Source: Elaboration of DataBench Survey, October 2018 (% of respondents, multiple answers)

5.3. Correlation Analysis, Approach and Results

This section analyses the DataBench survey by means of statistical correlation analysis techniques. This approach aims to analyse survey data to investigate the relationship between business and technical choices in BDT application projects.

The input to this analysis is a subset of questions focusing on business and technical performance. Questions concerning geographical location, company size, sector and use case have not been considered in our correlation models, as they would introduce a high number of variables and increase the risk of statistical overfitting and a consequent lack of statistical significance.

The analysis is carried out by applying factor analysis. Conceptually, factor analysis is a statistical tool that aims to group observed variables into higher-level constructs that have an intuitive meaning and can identify general and high-level rules governing the phenomena observed. This technique is also useful when dealing with data sets where there are numerous observed variables that can be hypothesised to reflect a smaller number of underlying/latent variables (Brown, 2014).

5.3.1 Methodology

The analysis was carried out using the psych package in R (Revelle). As a first step, the data set is transformed into numerical format. In doing so, questions that enable multiple answers — for example, KPI rating measured through a semantic scale (not at all important, slightly important, moderately important, very important, extremely important) — are transformed into a new format including *<question, answer, rate>*. Questions with a single possible answer containing a hierarchical structure, such as data size, are transformed into a new tuple (*<question, answer, level>*) where the third element of the tuple is a value on a likert scale, embeds a ranking and can be analysed with statistical techniques. Each value of the likert scale is mapped onto a numerical label preserving the hierarchy of the likert scale.

Questions with multiple possible answers not containing a hierarchy are transformed by using one-hot encoding. This technique replaces the question with a set of tuples (<question, answer, value>) where value is 1 if that particular answer was given, or 0 otherwise.

After the transformation, the correlation analysis is performed using a polychoric technique. Polychoric correlation estimates the correlation between two theorised normally distributed continuous latent variables, from two observed ordinal variables. This correlation is the most suitable for survey data with rating scales with a small number of response options (e.g., not important to extremely important). The polychoric correlation coefficient is between -1 and 1, where 0 identifies no relationship and 1 identifies a perfect relationship (positive or negative).

5.3.2 Results

Figure 30 shows the correlation matrix expressed by using a colour scale, where each cell identifies the correlation between the corresponding question-answer reported on the axes.

The analysis stresses a weak positive correlation, represented in light red, between answers related to the same question stressing the co-occurrence of all the options given. This behaviour can be identified in questions regarding business performance:

- Business KPIs for measuring the impact of Big Data and analytics efforts (cost reduction, time efficiency, product/service quality, revenue growth, customer satisfaction, business model innovation and increase in the number of new products and services launched) all appear to be positively correlated to each other, expressing an ambition to get multiple business benefits.
- The measured business impact is also positively correlated with an expected improvement on the same set of KPIs, as expected (confirmation by triangulation).

A similar positive correlation can be identified between expected and measured business benefits and technology investment.

Moreover, the analysis stresses a negative weak correlation between:

- Status of Big Data usage and expected business benefits, and
- Level of investments in BDTs and the impact of Big Data adoption on actual business benefits

Similarly, the analysis reveals a negative correlation between:

- Type of analytics and expected business benefits
- Type of analytics and level of investments in BDTs

The type of analytics is classified as descriptive, diagnostic, predictive and prescriptive. A negative correlation indicates that a greater orientation towards descriptive and diagnostic applications of BDTs is correlated with greater expected benefits and planned investments. The type of analytics shows a positive correlation with data size and real-time integration into business processes.

These correlations confirm the existence of a relationship between business and technical variables. They also provide consistent results triangulating questions that are repeated in the questionnaire from slightly different perspectives (e.g., types of expected benefits and their actual measure). Correlations are also consistent with descriptive analyses (see Section 5.2.2).

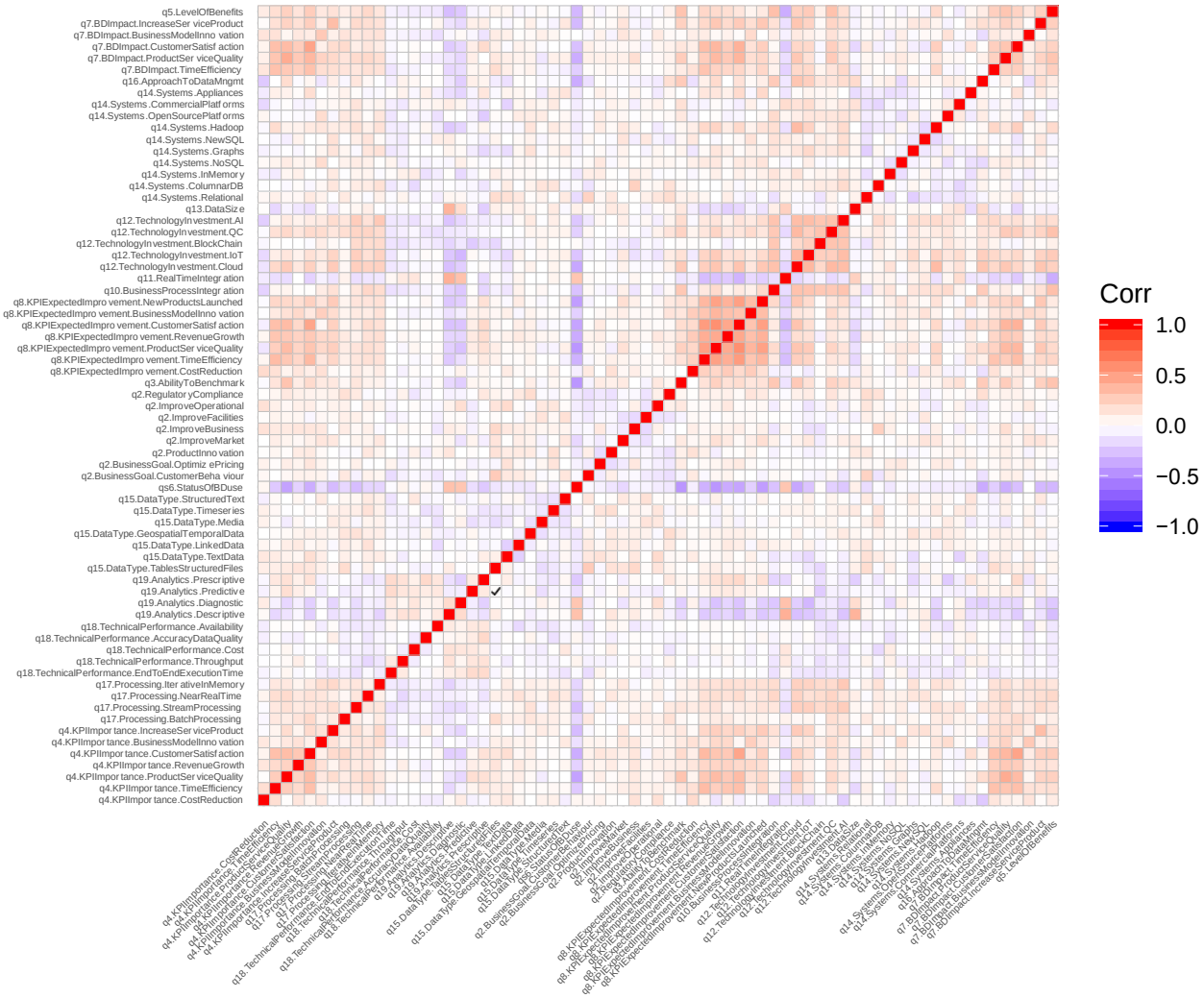


Figure 30 — Correlation Matrix

Source: Elaboration of DataBench Survey, October 2018

To aggregate correlated variables into a unified high-level model, correlations are further investigated using factor analysis. Factor analysis is a statistical method used to describe variability among observed, correlated variables by using a lower number of unobserved latent variables called factors. The observed variables are modelled as linear combinations of the latent factors. The underlying assumption of factor analysis is that the variation of a set of observed variables reflects the variation in a smaller set of unobserved variables. Factor analysis searches for such joint variations in response to unobserved latent variables. Factor analysis is typically recognised as potentially relevant whenever correlations involve multiple variables in the same conceptual direction and with multiple self-consistent findings.

Before factor analysis it is necessary to discard observed variables that are not useful to the process and to choose the number of latent factors to be modelled.

Observed variables that do not contribute to the analysis are discarded by using the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy, as a measure of factorability, to discard.

The KMO assesses if the partial correlations within the data set are close enough to zero to suggest the existence of at least one latent factor. By inspecting the KMO of each variable it is possible to drop the variables that provide too small a contribution, where the minimum acceptable value is 0.4. In our context this method discards variables regarding data storage and analytics environment as well as data types and business goals driving the adoption of Big Data.

The number of statistically significant latent factors is identified by using the scree test. The scree test represents the eigenvalues in a downward curve by ordering them from the largest to the smallest. Eigenvalues identify the amount of variance that can be explained by a factor, so they are useful in determining the number of factors to be extracted. The scree test assesses the number of significant factors according to the "elbow" of the graph where the eigenvalues seem to level off. Factors on the left of this point are considered significant. According to Figure 31, factor analysis models **five latent factors**.

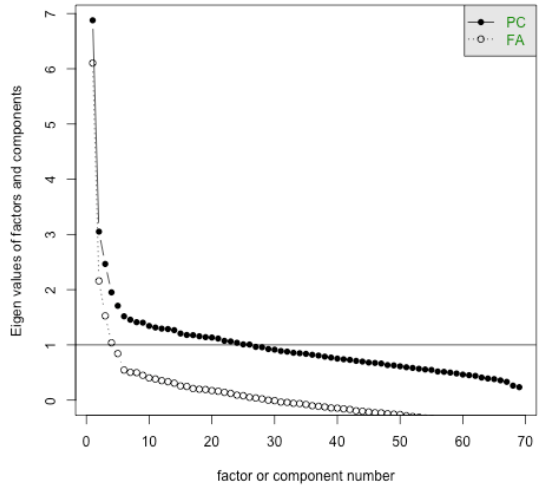


Figure 31 — Scree Plot of the Factor Analysis Models Five Latent Factors

Source: Elaboration of DataBench Survey, October 2018

Results of the factor analysis are reported in Table 7. Each cell represents the factor loading with respect to a variable. Factor loading is similar to correlation coefficients and varies from -1 to 1 — the closer loadings are to 1 the more the factor affects the variable. Variables with loadings smaller than 0.3 are discarded. A negative loading expresses a negative impact of the factor with respect to the variable, whereas a positive correlation expresses a positive impact.

	F1	F2	F3	F4	F5
q4.KPIImportance.TimeEfficiency			0.369		
q4.KPIImportance.ProductServiceQuality			0.402		0.303
q4.KPIImportance.RevenueGrowth			0.326		
q4.KPIImportance.CustomerSatisfaction			0.572		
q17.Processing.BatchProcessing			0.397		
q17.Processing.StreamProcessing		0.300			

	F1	F2	F3	F4	F5
q17.Processing.IterativeInMemory		0.339			
q19.Analytics.Descriptive				0.470	
q19.Analytics.Diagnostic				0.310	
qs6.StatusOfBDuse					-0.377
q3.AbilityToBenchmark					0.453
q8.KPIExpectedImprovement.CostReduction	0.449				
q8.KPIExpectedImprovement.TimeEfficiency	0.452				
q8.KPIExpectedImprovement.ProductServiceQuality	0.562				
q8.KPIExpectedImprovement.RevenueGrowth	0.532				
q8.KPIExpectedImprovement.CustomerSatisfaction	0.499				
q8.KPIExpectedImprovement.BusinessModelInnovation	0.600				
q8.KPIExpectedImprovement.NewProductsLaunched	0.576				
q10.BusinessProcessIntegration		0.363			
q11.RealTimeIntegration				0.617	
q12.TechnologyInvestment.Cloud		0.401			
q12.TechnologyInvestment.IoT		0.500			
q12.TechnologyInvestment.BlockChain		0.439			
q12.TechnologyInvestment.QC		0.358			
q12.TechnologyInvestment.AI		0.568			
q13.DataSize				0.380	
q16.ApproachToDataMngmt			-0.351	0.376	
q7.BDImpact.TimeEfficiency		0.320			
q7.BDImpact.ProductServiceQuality		0.407			
q7.BDImpact.CustomerSatisfaction		0.539			
q7.BDImpact.BusinessModelInnovation			0.351		

Table 7 — Factor Analysis Results

Source: DataBench Survey, October 2018, 700 interviews

Moreover, the method provides loadings between latent factors useful to express relationships between latent factors. Loadings weaker than 0.3 are discarded. In particular, F1 and F4 are related by a negative 0.3 loading, and F1 and F3 by a 0.5 positive loading. F5 and F2 are related by a 0.3 loading and F4 and F3 are related by a negative 0.3. Figure 32 represents the analysis to a higher level of abstraction.

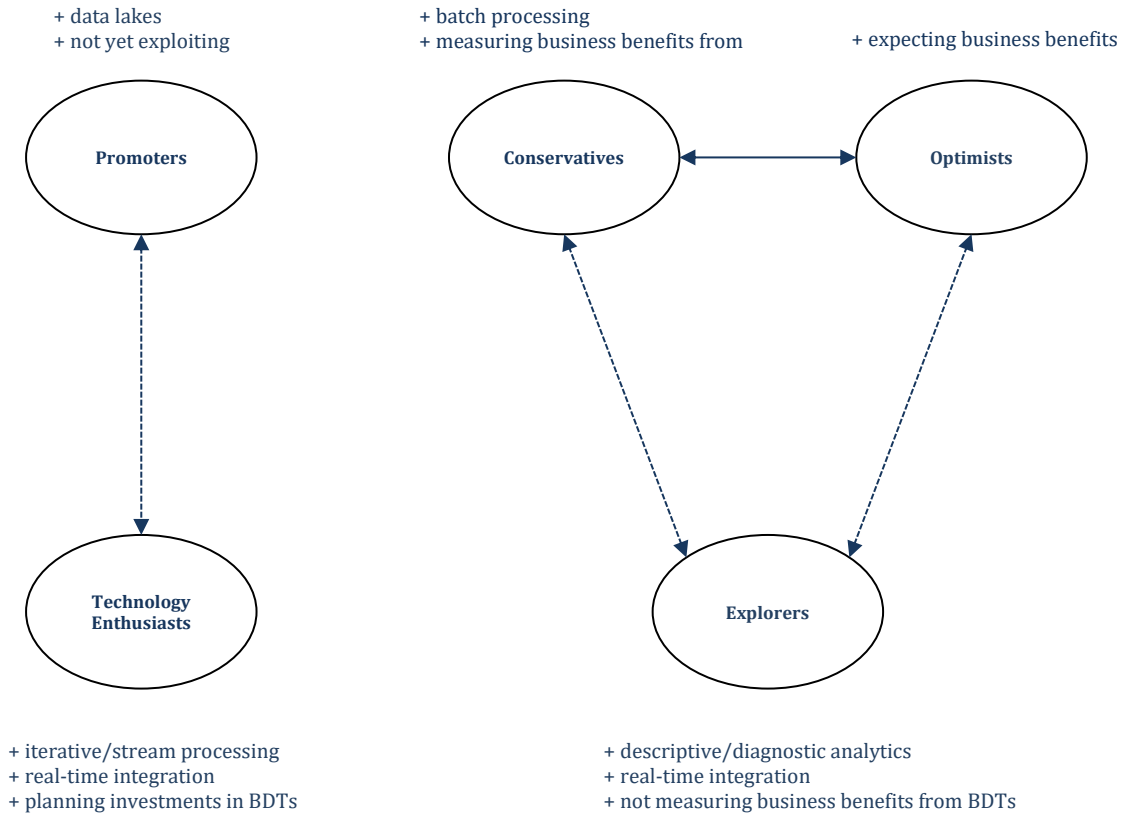


Figure 32 — High-Level Representation of Latent Factors

Source: Elaboration of DataBench Survey, October 2018

In the analysis, **four out of five latent factors can be interpreted as an approach towards the use of BDTs and their expected or actual business benefits**. One factor, factor F5, models expected business benefits but lacks a reference to technical choices. Each latent factor can be seen as a group of respondents, to which we have provided an intuitive label in Figure 32, namely Promoters, Technology Enthusiasts, Explorers, Conservatives and Optimists. A loading (correlation) between these groups indicates a significant overlap among respondents.

The five groups of respondents show a different level of maturity in the adoption of BDTs with different levels of integration of BDTs in their business processes:

- Promoters: From a business perspective, they are used to benchmarking business benefits. From a technical perspective, they have already adopted a data lake to store and manage their data, but so far they have not exploited Big Data.
- Technology Enthusiasts: They are a group of respondents with an innovative data-driven approach to BDTs. Technology Enthusiasts are currently planning investments in BDTs and believe in real-time integration with business processes to improve Big Data impact (in the future). These respondents aim to focus on iterative and streaming processing paradigms, but they have not yet measured benefits from BDTs.

Promoters and Technology Enthusiasts are positively correlated with each other, indicating that companies sharing the attitude of promoters (that is, used to measuring business benefits and infrastructurally ready with a data lake) are likely to be technology enthusiasts

and plan to be at the forefront of innovation with real-time, streaming data processing and a high level of integration of BDTs in their business processes.

- Conservatives: These companies successfully measure their business benefits and assess the impact of their BDT choices. From a technical standpoint, they mostly rely on batch processing of traditional relational data. They expect significant business benefits in the future, as emphasised by the positive correlation with the group of optimists.
- Explorers: These companies are highly orientated to data analytics, specifically to descriptive and diagnostic analyses, and to real-time integration of business processes. So far, this group appears to be focused on innovation and is not measuring business benefits, nor is it expecting business benefits in the near future. This attitude is seen in the negative correlation between Explorers and Optimists.
- Optimists: These companies believe they will obtain benefits from the application of BDTs in the future.

The contrast between the approach adopted by Explorers and Conservatives is emphasised by the negative correlation between these two groups. This emphasises the challenges raised by real-time integration and analytics to traditional Big Data users and the need for Explorers to move from innovation to business impacts.

To conclude, the factor analysis highlights that:

- Companies that have already obtained and measured business benefits from BDT projects are focused on traditional batch processing.
- Companies that experiment with more advanced real-time applications of BDTs have not yet measured business benefits.
- Companies that have not yet exploited BDTs or have a traditional exploitation of BDTs (batch) are technology enthusiasts and/or plan to explore more innovative applications of BDTs, but do not view future business benefits as measurable with economic KPIs at this stage of development of BDTs.[4]

5.4. Concluding Remarks

In conclusion, we can highlight a number of interesting features in the relationship between technical and business metrics, KPIs and the consequent choices.

Analysing the increasing complexity of the Big Data environment, with the ever-increasing pace of data velocity and the broadening of data types (variety), highlights the need to integrate heterogenous data.

From this perspective it is interesting to note the upcoming trend for real-time data analysis and technological solutions. The differing data types, sources and velocities at which data is streamed or extracted from storage are shaping the adoption of faster and more reliable real-time processing paradigms and data management approaches.

4 Brown, Timothy A. Confirmatory Factor Analysis for Applied Research. Guilford Publications, 2014. Revelle, W. (2018) psych: Procedures for Personality and Psychological Research, Northwestern University, Evanston, Illinois, USA, https://CRAN.R-project.org/package=psych Version = 1.8.12

Within this heterogeneous context of data management, new shapes of processing paradigms are being adopted. There is an opportunity for predictive and prescriptive analytics to better address the changing and challenging demand for improvements in processes and operations and accuracy of forecasts. Predictive and prescriptive analytics will ease the tasks of European companies in achieving increases in revenues and reduction in cost through the delivery of optimised products and services, improvements in customer satisfaction and optimisation of processes. IDC research indicates that although most European organisations are currently using simpler analytics techniques — 58% of European companies are efficiently using descriptive analytics — the near future looks brighter. The number of organisations interested in adopting more advanced analytics techniques is growing, with 43% considering use of predictive analytics and 41% showing an interest in prescriptive analytics.[5]

This qualitative analysis confirms the first takeaway provided by the statistical analysis: that, overall, European organisations are seeking to achieve a multiple set of business benefits, rather than focusing efforts and investments on improving one benefit at a time. However, as noted in 2.3.2, there are some exceptions if anticipated KPI impacts from BDA are analysed on a sector-by-sector basis.

A second line of outcomes comes from the statistical analysis, which is strictly about the correlation between business interests and technical needs.

It is interesting to consider that Big Data usage and the *expected* business benefits have a negative weak correlation. This means that companies taking most advantage from Big Data analytics and technologies to obtain insights initially expect a reduction, or at least a stagnation, of business benefits.

It is also worth noting the negative weak correlation between investments and *actual* business benefits — that is, increasing the investments in BDTs doesn't impact, or slightly negatively impacts, actual business benefits. It appears that a significant number of users of, and investors in, BDTs are nevertheless sceptical about achieving significant business benefits in the near term.

This may be contextualised within an uncertain environment, where data is perceived as scarce by most European companies (consider that the fabric of the European economy is characterised mostly by SMEs) and/or companies are not confident in BDT choices due to scarce knowledge of the technologies and the state-of-the-art solutions tailored to their own context.

Another interesting point concerns the selection and deployment of specific types of data analytics (descriptive, diagnostic, predictive, prescriptive, etc.). The insights on data analytics show a negative weak correlation with two objects of analysis: *expected* business benefits and level of investments. In the first case, the correlation is negative weak, meaning that when shifting from a more traditional data analytics solution (descriptive) to a more modern and complex solution (prescriptive) companies perceive a reduction in the expected business benefits. This is possibly due to the lack of confidence in the data set owned or, more probably, to limited confidence in the potential of more advanced and often obscure data analytics techniques. In the second case the correlation between the two

5 How European Organisations Are Benefitting from Big Data and Analytics (IDC #EMEA44880219, March 2019)

variables is negative weak, meaning that European organisations are more willing to invest more money in simple and "old-fashioned" analytics techniques (descriptive and diagnostic) rather than investing in brand new and more powerful techniques (predictive and prescriptive).

There are several possible explanations for this.

In the first scenario, the increasing complexity of the analytics also increases the cost of the technology, so companies are not ready to invest, especially if they believe they do not have access to a sufficient amount of data and/or have poor data quality.

In the second scenario, it is possible to claim that as the European market is dominated by SMEs, such small companies are not yet ready (or not confident enough) to tackle Big Data challenges using modern and effective BDTs.

A last consideration is driven by the identification of five categories of respondents, based on factor analysis. By identifying the different stages of user maturity, it is possible to drive insight into the interaction of players in the Big Data environment.

This analysis also sheds light on the potentially puzzling negative correlations between use and investment in advanced BDTs and anticipated business benefit.

1) Conservatives are successfully using and benchmarking traditional BDA and expect to adopt and benchmark more advanced BDTs in the future.

2) Optimists expect to obtain benefits from BDA adoption in the future, but are not necessarily currently measuring BDA business outcomes. Nevertheless, there is a significant overlap with Conservatives.

3) Promoters are used to benchmarking the business benefits of IT investment and already have advanced BDTs in place (such as data lakes) but are not yet using and exploiting the business potential of this advanced infrastructure.

4) Explorers embrace the new challenges posed by the increasing complexity of the BDA environment, but are not currently focused on measuring, or even immediately achieving, business outcomes from BDTs. They adopt new BDTs for exploratory purposes, to facilitate the development of future BDA applications.

5) Technology Enthusiasts have adopted a data-driven philosophy and are adopting or seriously considering adopting real-time integration of BDTs into business processes.

These categories, and their importance for maximising the economic impact of BDA, are further explored in the Conclusions of this report.

6. General Conclusions

6.1. Overview

The overall objective of DataBench is to provide benchmarks and performance evaluation mechanisms to identify the business impact and the industrial significance of the deployment of Big Data technologies in Europe.

Successful deployment of BDTs in the medium to long term depends on adopters' ability to convincingly benchmark and assess the impact on their businesses. This will primarily be based on adopters' own business objectives, but also to enable comparison with and learning from best practice in their sector.

Continued assessment of business contribution is necessary not only to justify continued financial support for and investment in BDTs within adopter organisations, but also to enable BDA deployments to be tuned and improved to continue to make the most effective contribution as the company, and its external market and business environment, change over time.[6]

Accordingly, the DataBench survey asked respondent companies not only how important they felt such benchmarking to be, but which business KPIs they expected to be more or less significantly improved by their existing or planned BDA investment. 45% of respondents considered benchmarking the business impact of BDA to be very or extremely important.

Although there was a generally broad consensus on the KPIs that respondents considered would be most beneficially impacted by BDA deployment, there were some significant variations by industry sector and business size. As a step towards helping companies to design their own internal and sector comparative business benchmarking strategies, we have further analysed these differences in KPI emphasis for this report.

We also further analysed the technology approaches that respondents were currently using, and were planning to use in future, for BDA to identify correlations between these and their expectations of the contributions that BDA would make in business terms.

6.2. User Needs by Industry

As reported in the previous D2.2 DataBench report (Preliminary Benchmarks of Industrial Significance), sectors are at different stages of BDA adoption.

6 IDC research suggests that BDA "inward" analysis is important but "outward" understanding is a key factor for future success. Engaging in benchmarking against BDA best-in-class is at least as important as simply evaluating internal benefits of BDA adoption. (How do European Organisations Measure BDA Performance? IDC #EMEA44834419, February 2019)

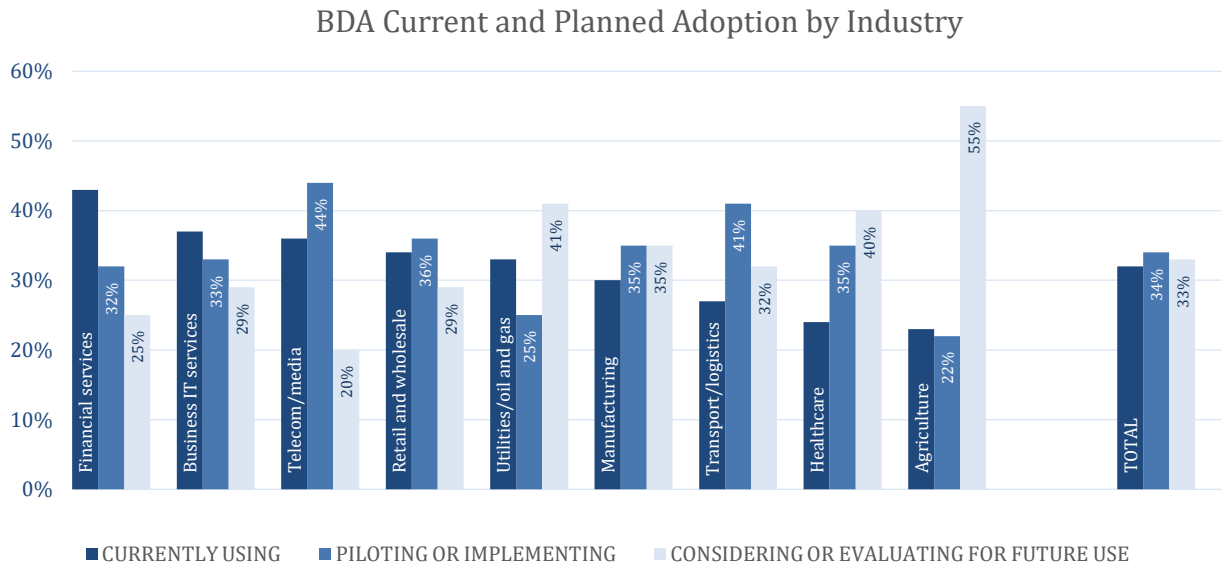


Figure 33 — Current and Planned BDA Adoption by Industry (% of respondents)

Source: DataBench Survey, October 2018

To establish best practice in benchmarking business impact on a per-sector basis, in this report we extended the analysis to identify the sectors that had significantly higher or lower expectations of BDA to have a positive business impact on specific KPIs within their organisations. This identified not only a significant difference in KPI expectations between many of the sectors, but enabled sector approaches to be informally categorised.

	Category
Agriculture	Conservative
Healthcare	Conservative
Retail and wholesale	Conservative; seeking transformation
Transport/logistics	Conservative; seeking transformation
Manufacturing	Broad positive expectations, including transformation
Financial services	Clear main focus
Utilities/oil and gas	Clear main focus
Business/IT services	Aspirational for both operational focus and transformational KPIs
Telecom/media	Aspirational for both operational focus and transformational KPIs

Table 8 — Categorisation of Distinctive KPI Priorities by Sector

Source: IDC Elaboration of DataBench Survey, October 2018

This is broadly consistent with the current and planned adoption chart above, but with additional insights into sector attitudes towards BDA, including:

- Although agriculture (particularly) and healthcare, the two sectors with the lowest levels of current BDA adoption, both indicate high levels of considering or evaluating future BDA adoption, their actual expectations of the importance of BDA in contributing to the seven KPIs are significantly below other sectors. This is disappointing given the significant contributions that BDA could make in both sectors and suggests that BDT suppliers have not yet convinced these sectors of the potential value of BDA.

- Retail and wholesale and transport and logistics have significant expectations for digital transformation arising from BDA (the business model innovation KPI), and this is considered significantly more important to them than incremental improvements in core KPIs.
- Although most sectors have a broad range of expectations for BDA in terms of KPIs (see para 5.3.2), financial services and utilities/oil and gas both have a significant focus on KPIs — customer satisfaction and revenue growth, respectively.

These sector-specific KPI priorities should be considered by individual companies when assessing their own benchmarking strategy for BDA impact, and also taken into account by BDT vendors seeking to address sector priorities.

We also compared the most-often-cited primary BDA use cases in each sector with the KPIs deemed to be most important for measuring BDA business impact. There was a good general correlation between the business benefits that would be expected from the most popular use cases and the stated KPIs, with the slight exception of the healthcare sector, where several of the top use cases would be most expected to deliver efficiency increases and associated cost reduction.

This highlights, however, the somewhat surprising finding that although financial services, manufacturing and transport/logistics considered cost reduction more significant as a BDA KPI than other sectors, overall it is by far the least-favoured KPI, in all sectors, in terms of assessing BDA impact when compared with the other six KPIs in the DataBench model.

Many use cases, for example in BDA supply chain applications, have significant potential for cost saving. This may suggest that BDA adoption, especially if significant investment is required, is more easily justified internally to a company by appealing to its ability to boost less traditional KPIs. It also suggests, perhaps, that opportunities to deploy modern BDA to effectively address less exciting but equally important business goals such as cost efficiency are not always being fully addressed by adopters.

Analysis of the Horizon 2020 BDA-related projects (Chapter 4) shows that the eight vertical projects are mostly pursuing development of BDTs in line with the business needs clustered by industry sector, both from a KPI and use case perspective.

Five of the eight vertical projects analysed present objectives consistent with the needs of the vertical in which they are classified. This perfect match between needs of the vertical and project scope indicates a clear understanding of the context in which the projects are operating to develop industry-specific BDA and BDTs. Only three of the vertical Horizon 2020 projects analysed have a mismatch between the industry needs identified with the KPI and use case analysis in Chapter 2 and the type of BDA solutions being developed. However, this mismatch shouldn't suggest that the projects aren't effectively contributing to the improvement of the sector's needs: first, within the DataBench survey, not all possible sector-specific use cases could have been investigated, and some of the projects' objectives can be evaluated with more niche sector-specific KPIs.

6.3. Business Size Factors and SME Needs

As reported in the previous DataBench deliverable (D2.2), the level of current and planned BDA adoption varies significantly by company size.

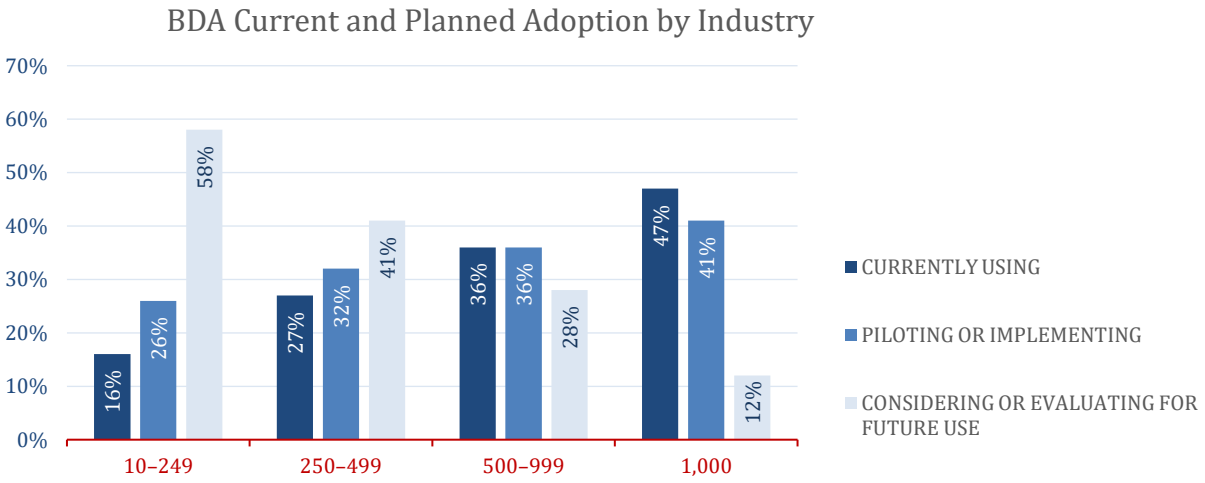


Figure 34 — Current and Planned BDA Adoption by Industry (% of respondents)

Source: DataBench Survey, October 2018

The most notable difference is between SMEs, which significantly lag medium-sized and large enterprises in terms of adoption of BDA (with only 16% of them already using BDA), and large companies with over 1,000 employees (47% of which already use BDA). Given the vital role that SMEs play in the European economy, the low adoption of BDA must be a cause for concern.

We further investigated the prevailing attitudes towards BDA adoption by different sized businesses in Chapter 3. The survey results indicate that cost reduction is significantly more regarded as a significant KPI for benchmarking BDA benefits by SMEs than for other company sizes. All other KPIs, except time efficiency, are less regarded by SMEs as important to assess BDA impact than by larger companies, with the product/service quality, revenue growth, customer satisfaction and business model innovation KPIs being considered significantly less important.

It appears that the surveyed SMEs do not consider it likely that they will deploy BDA to address many of the key business KPIs that will underlie success and growth, and that they will concentrate on using BDA to achieve cost savings. While it is understandable that small companies must address cash flow/burn rate as a priority, there are numerous other valuable applications of BDA that could be of significant benefit, such as improving their understanding of the needs of their target market and making informed decisions about how best to meet them.

In sectors such as agriculture there are also very significant opportunities to take advantage of new technologies to improve effectiveness — using IoT and remote sensing data for smart farming and precision agriculture, for example.

Small companies face many challenges in BDA adoption, of course, including:

- Cost of adoption
- Lack of in-house data on which to base analysis
- Lack of in-house analytical skills

However, there is a significant opportunity for suppliers of BDT and BDA solutions to package these in ways that address these obstacles, to supply affordable, out-of-the-box analytics-based solutions that can be used without specialised skills and that draw on

external data sources — for example, environmental data or market research data — in ways that are relevant to SMEs' business needs.

This is already taking place, to an extent. The Big Value Data Association (BVDA), for example, identifies companies, themselves SMEs, that are already packaging analytics solutions for agriculture and healthcare.[7]

The survey results indicate, however, that the SME sector itself does not appear to see BDA as important to other KPIs as to the understandable but relatively unambitious goal of cost reduction.

Initiatives to encourage the BDA vendor community to more effectively address the market opportunity for SME BDA adoption, and to encourage a similar market pull from SMEs for usable and relevant analytics solutions, would potentially have significant economic benefit.

6.4. Types of BDA Adopter and technology Choices

Qualitative and quantitative analysis of potential correlations between survey respondents' business expectations from BDA adoption and their current and planned analytics technical infrastructure is detailed in Chapter 5.

In the initial findings, we were surprised to see a slightly negative correlation between the current status of BD usage by adopters and the expected business benefits, and the level of investments in BDTs and the impact of BDA adoption on actual business benefits.

These, however, can be explained by a further statistical factor analysis of the survey results to identify patterns of BDA adoption among respondents.

Five (to some extent overlapping) categories of BDA adopter were identified based on common patterns in their survey responses:

1) Conservatives are already using more traditional tools and solutions for BDA (for example, relying on batch processing of traditional relational data), are effectively benchmarking existing BDA business benefits, and expect to implement and benchmark significant business benefits from more ambitious BDT adoption in future.

2) Optimists believe they will achieve and measure business benefits from further adoption of BDTs (so there is significant overlap with the Conservatives) but are not necessarily already assessing and benchmarking business benefits from their current BDA usage.

3) Promoters are used to benchmarking business benefits from IT, including existing traditional BDA, and have already put advanced BDTs in place (data lakes, for example) but are not yet exploiting and measuring the potential of such advanced BDTs.

4) Explorers are investigating real-time integration solutions and engaging in both traditional and more advanced analytics solutions (such as diagnostic as well as descriptive analytics). These companies are currently not focused on measuring or even necessarily achieving business outcomes from BDTs: they are adopting new BDA solutions and tools for exploratory purposes to enable fast-track implementation of business use cases yet to be

7 SMEs in the European Data Economy (http://www.big-data-value.eu/wp-content/uploads/2018/01/SMEs-Brochure-2017.pdf)

identified. There is therefore a negative correlation with the Optimists, for which benchmarking of anticipated and delivered business outcomes is an expected priority.

5) Technology Enthusiasts strongly believe in the potential of the insights driven by BDA and are preparing for or already adopting the most advanced BDTs (such as real-time process integration and iterative/stream processing). However, in common with Explorers, they have not yet measured business benefits from BDTs, nor is this an early priority.

This analysis reveals that, as with business intelligence adoption over past decades, there is a range of motivations — from purely pragmatic business-driven adoption strategies, often through standalone proof-of-concept projects, to technology-led innovation to facilitate the implementation of business-led use cases at a later stage.

This explains the apparent negative correlation between adoption of and investment in advanced BDTs and the expected business benefit. The underlying cause is the proportion of adopters that are in the Promoter, Explorer and Technology Enthusiast categories and are not yet even trying to formally measure business benefits of BDA adoption — and may well expect these benefits not to arise for some time after enabling BDT infrastructure has been put in place.

However, even for these BDT pioneers, there will come a time when, even in the largest and most sophisticated IT-user companies, their investment in BDA will need to demonstrate convincing business benefit. The goal of DataBench to converge technical and business perspectives on BDA adoption is just as important for such companies, if not more so, than those in the Conservative and Optimist categories, which will seek guidance on benchmarking best practice from the start.

A fruitful area for further analysis would be to consider the distribution of the adopter categories by business size, particularly, and by sector.

6.5. Considerations About Future BDA Developments

The BDA field is in a constant state of innovation. Although a significant proportion of the DataBench survey respondents are still largely in the planning stage in terms of using today's BDTs, even those well advanced are looking to benefit from ongoing improvements across the field. Examples include:

- Greatly improved user interface support for BDA users, with specific support for analytics experts and non-technical business analysts to cooperate effectively to produce and maintain ever more effective BDA outcomes
- Continuously improving data exploration capabilities, enabling experts and non-experts to locate, understand and analyse data from within and outside their organisations in ways and on timescales that previously would have been impractical and with automated assistance to identify potentially significant patterns and features in such data
- Ever more capable underlying hardware and software technologies making ad hoc analysis of huge data sets from different sources practical and enabling "what if" experimentation to be carried out in hours rather than months

And, of course, the opportunity to take advantage of AI and machine-learning technologies. IDC has described the use of AI as "a slow-motion explosion" with significant opportunities and risks in the coming years. IDC's 2018 AI User Survey concluded that, "The adoption of AI technologies and solutions is still relatively low in Europe — around 14% of

organisations today — but it is increasing quickly. This rapid growth is set to continue in the coming years."[8]

Benchmarking the business impact of BDTs, and technical benchmarking of BDT capabilities to support it, needs to take account of these new and expanding technologies and application areas and identify best practice, high-value use cases, and obstacles and opportunities in terms of the technical infrastructures needed to implement it.

6.6. Next Steps

The next step in WP2 is to launch a second wave of the DataBench survey, inviting the industrial partners of the ICT 14-15 projects and other projects in the BDT field, interested in measuring KPIs, to compile the same questionnaire and compare their results with their peers through the self-assessment tool presented in Chapter 4. This will start in July 2019. The team hopes to involve a number of industrial partners from these projects to collect data to further refine and finalise the industrial KPIs.

The analysis carried out in this report will feed into the research carried out by WP4 on the evaluation of business performance in specific case studies, as also documented in the D.4.2 report delivered in parallel with this report. More importantly, the results of this analysis will be tested and validated through the case studies carried out by WP4 and this will lead to the finalisation of the benchmarks for D.2.4, due in December 2019.

7. Annex

7.1. References

IDC Worldwide Big Data and Analytics Software Forecast, 2018–2022 (IDC #US44243318, September 2018)

SMEs in the European Data Economy (http://www.big-data-value.eu/wp-content/uploads/2018/01/SMEs-Brochure-2017.pdf)

Dubé, L., & Paré, G. (2003): Rigor in Information Systems Positivist Case Research: Current Practices, Trends, and Recommendations/MIS Quarterly, 27(4), 597-636

Brannen, J. (Ed.). (2017): Mixing Methods: Qualitative and Quantitative Research/New York, NY: Routledge, 2017

Brown, Timothy A. Confirmatory Factor Analysis for Applied Research/Guilford Publications, 2014

Revelle, W. (2018) psych: Procedures for Personality and Psychological Research, Northwestern University, Evanston, Illinois, USA, https://CRAN.R-project.org/package=psych Version = 1.8.12

7.2. Glossary

The following glossary presents the definitions of the terms used in this study.

8 AI in Europe: Key Findings of IDC's 2018 AI User Survey (IDC #EMEA44220518, August 2018)

BIG DATA

According to IDC,[9] the definition of Big Data is based on the following main attributes:

1) The escalation of the dimension of the data sets ("volume")
2) The increasing pace at which data is produced and exchanged ("velocity")
3) The growing number of types of data to deal with alongside the increase in sources that produce data ("variety")

These attributes are measured, according to IDC, with the following parameters:

Volume:

- Data collected is over 100TB

Velocity:

- If high-speed messaging technology for real-time streaming data integration and analytics is at or above 60GBps, and/or
- If the data sets may not be very large today but are growing rapidly at a rate of 60%+ annually
- If technology is deployed on dynamically adaptable infrastructure

Variety:

- If the data sets include two or more data types or data sources (structured/unstructured data flows, numerical, text, audio, video, pictures, images and metadata)
- If the data sets include high-speed data sources, as clickstream tracking or monitoring of machine-generated data
- If different company data sources are employed — for example, data from procurement, clients' details from marketing (client segmentation and profiling) and from accounts (customer personal data), clients' purchases from sales, inventory data from logistics, and so on

BENCHMARKS

Benchmarks are the result of the assemblage and calculation of performance metrics to be used for comparative purposes. In this task we focus on quantitative indicators of business performance improvements achieved by the use of Big Data technologies that can be used as target or best performance benchmarks by other organisations. Benchmarks are categorised according to type of use case, business process and industry.

EUROPEAN ECONOMIC SIGNIFICANCE

To employ economic indicators such as value added, revenues and employment to estimate the economic "footprint" of the BDT market in Europe and to identify the industry sectors where the prospective BDT business benchmarks will generate the highest potential economic impact.

INDUSTRIAL SIGNIFICANCE

9 IDC's Worldwide Big Data and Analytics Software Taxonomy, 2017 (IDC #US42353216)

To identify the BDT business benchmarks responding to the actual and emerging needs of industrial users, with the highest potential impact on business processes.

BIG DATA USE CASES

A discretely funded effort designed to accomplish a particular business goal or objective through the application of Big Data technology to particular business processes and/or application domains, employing line-of-business and IT resources. Examples are price optimisation, fraud risk assessment and customer profiling.

BUSINESS INDICATORS

In the DataBench indicators ecosystem,[10] business features correspond to the main parameters used to identify and classify the typologies of BDT implementations in a business organisation (use cases) and the performance metrics used to measure their business impacts (business KPIs). As shown in the figure below, the project conceptual framework aims to connect the technical benchmarking metrics with the main business impacts measured by business KPIs selected by this project.

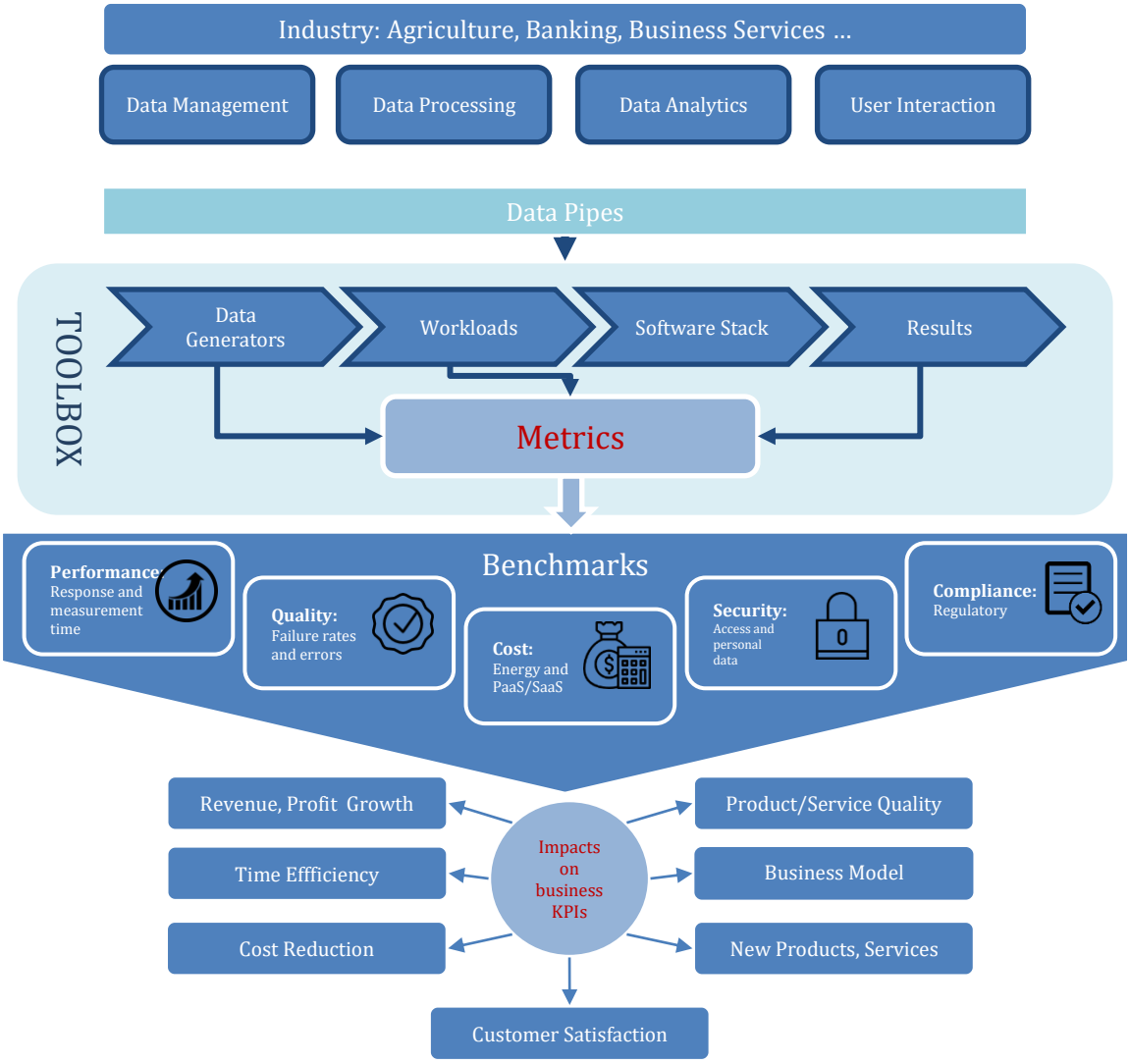


Figure 35 — Technical and Business Benchmarking Framework

Source: DataBench October 2018

The business features indicators can be divided into the following main groups:

1. Classification of business users (industry and company size)
2. Type of BDA implementation (application area, level of business process integration, level of BDA solutions maturity, company approach to data management, main business goals)

10 DataBench, D.1.1: Industry Requirements, Benchmarking Metrics and KPIs

3. Type of use case (cross-industry and industry-specific)
4. Business impact KPIs

The indicators categories are presented in detail in Table 9 and Table 10 below. Groups 1, 2 and 3 are semantic indicators measured through simple nominal questions (business users select the category in which they belong) to classify users. The survey results are measured as frequencies of respondents by category. Descriptive parameters can be used to measure the correlation between type of user and type of application and in turn type of business impact. They will be used in the benchmarking tool as a user interface to guide users to identify themselves and their type of BDA application, and in turn to look for the type of technical benchmark most relevant for them.

The use cases (group 3, presented in detail in Table 11 and 12) represent the link between technical solutions and business goals. The potential list is long, with a long tail of specific use cases. For survey purposes we have selected 12 cross-industry use cases and 23 industry-specific use cases, representing the most frequent and potentially impactful typologies identified so far by IDC research.

The business impact KPIs (group 4) are seven indicators selected on the basis of business literature and research as the most relevant for measuring innovative technology impacts (Table 10). They are measured as simple numeric values mainly in percentage (percentage of improvement).

Industry	Company Size	Application Area	Level of Business Process Integration
• Agriculture • Banking, insurance, other financial services • Business or professional services, excluding IT services • IT services • Healthcare • Manufacturing process • Manufacturing discrete • Retail trade • Wholesale trade • Telecommunications • Media • Transport and logistics • Utilities • Oil and gas	• 10 to 49 employees • 50 to 249 employees • 250 to 499 employees • 500 to 999 employees • 1,000 to 2,499 employees • 2,500 to 4,999 employees • 5,000 or more employees	• Customer service and support • Engineering • Research and development (R&D) • Product innovation (new business initiatives) • Maintenance and logistics • Marketing • Finance • HR and legal • Sales • Product management • Governance, risk and compliance • IT and data operations	• High (where there is real-time integration with business processes, e.g., for real-time fraud detection) • Medium (where there are mixed levels of integration with business processes, e.g., propensity models available as part of business processes but not scored in real time) • Low (e.g., where Big Data reports and dashboards are processed in a batch environment and made available the following day)

Table 9 — Business Parameters: Industry, Company Size, Application Area, Level of Business Process Integration

Source: DataBench Survey, October 2018

Level of BDA Solutions Maturity	Business KPI	Business Goals	Approach to Data Management
• Currently using • Piloting or implementing • Considering or evaluating for future use • No use and no plans	• Increase in the number of products/services launched • Customer satisfaction • Business model innovation • Revenue and profit growth • Product/service quality • Time efficiency • Cost reduction	• Better understanding customer behaviour and expectations • Optimise pricing strategies and go-to-market programmes • Product, services or programme improvement and innovation • Improve understanding of the market and competitors • Improve and optimise business processes and operations • Improve facilities, equipment design, maintenance and utilisation • Improve operational, fraud and risk management • Implement better regulatory compliance and financial controls	• Structured and transactional data is captured and processed in a data warehouse or operational data store • Structured and unstructured data from different locations, including on-premise and in the cloud, is captured and processed • Enterprisewide repositories or data lakes are used to capture, organise and process data from multiple sources and formats • Real-time data (e.g., log files, social media and IoT data) is streamed and used alongside other contextual data • The Big Data platform is available/exposed to external customers, partners and developers to build and extend data-driven benefits

Table 10 — Business Parameters: Maturity, Business KPIs, Business Goals, Approach to Data Management

Source: DataBench Survey, October 2018

Use Case	Industries
Price optimisation	All
New product development	All
Risk exposure assessment	All
Regulatory intelligence	All (excluding agriculture)
Customer profiling, targeting and optimisation of offers	Banking, insurance, other finance, business or professional services, IT services, retail trade, telecommunications, media, utilities
Customer scoring and/or churn mitigation	Banking, insurance, other finance, telecommunications, utilities
Fraud prevention and detection	Banking, insurance, other finance, business or professional services, IT services, healthcare, telecommunications
Product and service recommendation systems	Banking, insurance, other finance, business or professional services, IT services, retail trade, telecommunications, media
Automated customer service	Banking, insurance, other finance, business or professional services, IT services, healthcare, retail trade, telecommunications, media
Supply chain optimisation	Agriculture, manufacturing process and discrete, retail trade, wholesale trade, transport and logistics, utilities, oil and gas
Predictive maintenance	Agriculture, manufacturing process and discrete, wholesale trade, transport and logistics, utilities, oil and gas
Inventory and service parts optimisation	Agriculture, manufacturing process and discrete, wholesale trade, transport and logistics, oil and gas

Table 11 — Classification of BDA Cross-Industry Use Cases

Source: DataBench Survey, October 2018

Industry	Specific Use Cases	Industry	Specific Use Cases
Agriculture	• Precision agriculture • Yield monitoring and prediction • Field mapping and crop scouting • Heavy equipment utilisation	**Retail trade**	• Intelligent fulfilment
Banking	• Cyberthreat and detection	**Wholesale trade**	• Intelligent fulfilment • Increase productivity and efficiency of DCs/warehouses
Insurance	• Usage-based insurance	**Telecommunications**	• Network analytics and optimisation
Other financial services	• Cyberthreat and detection	**Media**	• Ad targeting • Scheduling optimisation
Business or professional services	• Social media analytics	**Transport and logistics**	• Connected vehicles optimisation • Logistics and package delivery management
Healthcare	• Illness/disease diagnosis and progression • Personalised treatment via comprehensive evaluation of health records • Patient admission and re-admission predictions • Quality of care optimisation	**Utilities**	• Field service optimisation • Energy consumption analysis and prediction
Manufacturing process	• Smart warehousing • Asset management • Quality management investigation	**Oil and gas**	• Energy consumption analysis and prediction • Field service optimisation
Manufacturing discrete	• Smart warehousing • Asset management • Quality management investigation • Connected vehicles optimisation		

Table 12 — Classification of BDA Industry-Specific Use Cases

Source: DataBench Survey, October 2018

7.3. The DataBench Survey Sample

The size of the sample was decided to ensure the statistical representativeness of the European industry with a margin of error as low as possible (3.5% for the whole sample) and to remain within the time and budget constraints of the project plan.

The survey used hard quotas by country/region and soft quotas by industry and size class to achieve a balanced sample of valid interviews.

The results displayed in the following chapters may include aggregated values for some subsegments to guarantee reliability or maximise explanatory value.

The list of countries surveyed was based on the following criteria:

- Geographical balance (representing all main geographical areas in the EU)
- Country size (mix of large, medium-sized and small Member States)
- IT maturity balance (mix of Member States with high, medium and low intensity IT spending)
- Share of data market value (the Member States selected represent 87% of the European data market value in 2017)[11]
- Adequate coverage of the EU economy (the Member States surveyed represent 76% of EU GDP in 2017)[12]

The survey results therefore are representative of EU-level results.

The final breakdown of the 700 interviews is shown in Figure 36 and Figure 37 and satisfies the targeted criteria, representing a balanced sample by country, industry and size class.

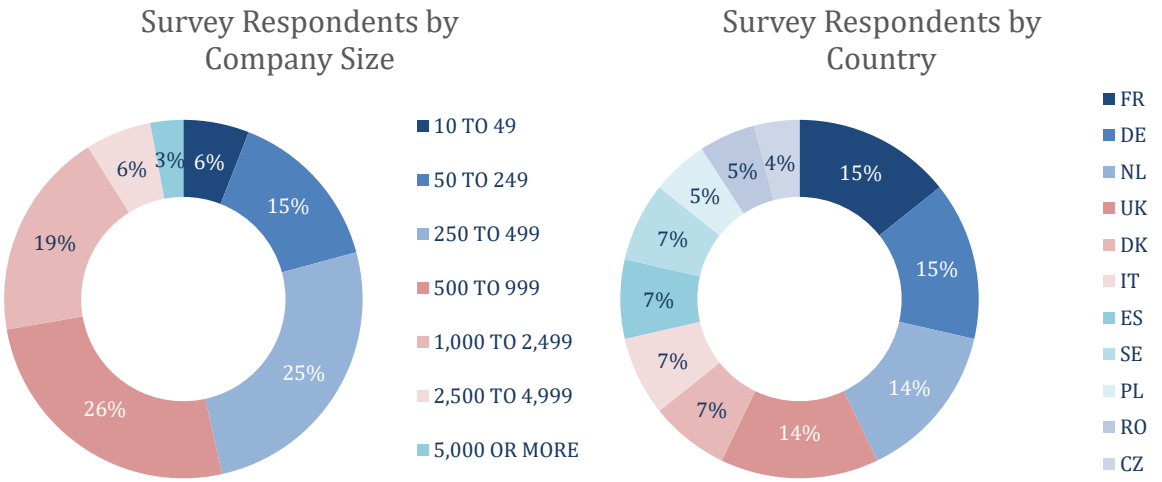


Figure 36 — Survey Sample by Size and Country

Source: DataBench Survey, October 2018, 700 interviews

11 Update of the European Data Market Study, Facts and Figures Report, January 2018 (IDC #US44243318)

12 Eurostat data, EIU, EC EU growth, December 2017

The steps involved in developing the survey questionnaire can be broadly categorised as follows:

I. Confirm survey objectives
II. Identify key research topics including:
    a. Importance of BDTs to business strategy and objectives
    b. Level of maturity in BDT adoption
    c. BDT planned use or current usage
    d. Use case priorities actual and planned
    e. KPIs and metrics used
    f. Perceived and actual business benefits and returns
III. Translate research topics into survey questions
IV. Organise and structure questionnaire into logical sections including:
    a. Introduction questions that define the objectives of the study and the sponsor/how the data will be used, incentives, confidentiality statement
    b. Survey screener or qualifying questions to identify target audience
    c. Warm-up key questions to identify where BDTs are being implemented, adoption drivers and so on
    d. Main topic questions as outlined in step 2 above
    e. Profiling demographics questions to enable post-survey data segmentation
    f. Questions to determine permission to use data and feedback on the survey
V. Peer review survey questionnaire
VI. Run quality checks to ensure the questionnaire meets best practice

The quality and level of completion of the interviews is good and corresponds to the targeted objectives.

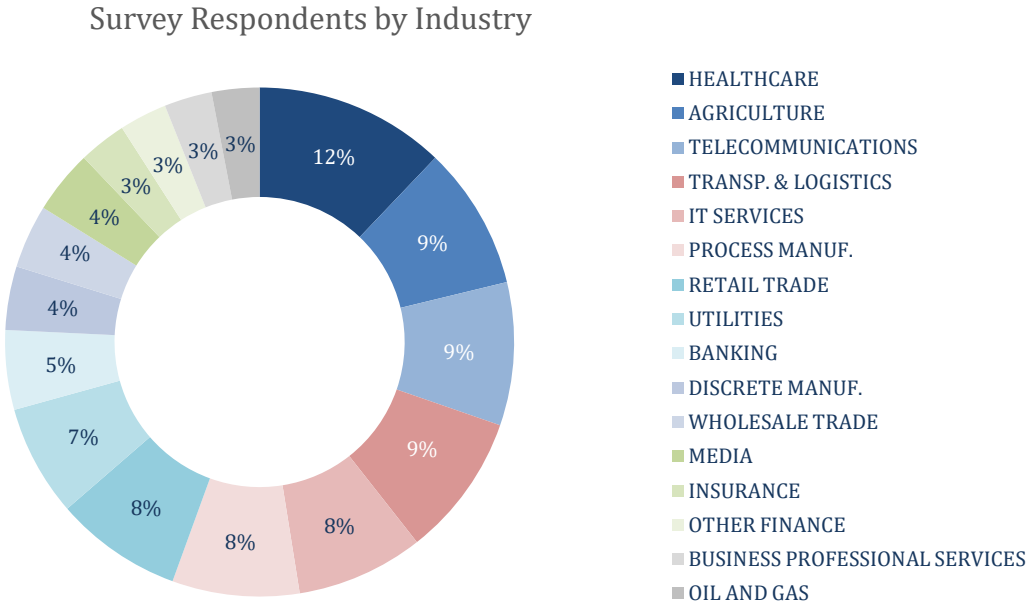


Figure 37 — Survey Sample by Industry

Source: DataBench Survey, October 2018, 700 interviews

7.4. BDA Use Cases by Industry

The tables below, sourced from the survey, present the list of use cases implemented by industry, ranked by the absolute number of responses and the percentage of respondents.

Agriculture	# Responses	% Responses	Finance	# Responses	% Responses
Field mapping and crop scouting	44	68%	**New product development**	56	73%
Price optimisation	42	65%	**Customer profiling, targeting and optimisation of offers**	55	71%
Inventory and service parts optimisation	42	65%	**Risk exposure assessment**	53	69%
Risk exposure assessment	38	58%	**Regulatory intelligence**	53	69%
New product development	37	57%	**Fraud prevention and detection**	50	65%
Predictive maintenance	37	57%	**Automated customer service**	47	61%
Yield monitoring and prediction	37	57%	**Price optimisation**	46	60%
Heavy equipment utilisation	37	57%	**Product and service recommendation systems**	45	58%
Supply chain optimisation	34	52%	**Customer scoring and/or churn mitigation**	43	56%
Precision agriculture	33	51%	**Cyberthreat and detection**	33	43%

Business/ IT services	# Responses	% Responses	Healthcare	# Responses	% Responses
Customer profiling, targeting and optimisation of offers	57	73%	**Regulatory intelligence**	52	67%
Risk exposure assessment	52	67%	**Fraud prevention and detection**	51	65%
Price optimisation	51	65%	**Quality of care optimisation**	51	65%
New product development	50	64%	**Automated customer service**	47	60%
Fraud prevention and detection	48	62%	**Risk exposure assessment**	46	59%
Product and service recommendation systems	47	60%	**Patient admission and re-admission predictions**	46	59%
Regulatory intelligence	46	59%	**Illness/disease diagnosis and progression**	44	56%
Automated customer service	41	53%	**Personalised treatment via comprehensive evaluation of health records**	44	56%
Social media analytics	13	17%	**Price optimisation**	43	55%
Customer profiling, targeting and optimisation of offers	57	73%	**New product development**	42	54%

Manufacturing	# Responses	% Responses	Retail and Wholesale	# Responses	% Responses
Regulatory intelligence	56	63%	**New product development**	54	65%
New product development	55	62%	**Price optimisation**	53	64%
Price optimisation	54	61%	**Supply chain optimisation**	49	59%
Supply chain optimisation	51	57%	**Intelligent fulfilment**	47	57%
Predictive maintenance	51	57%	**Regulatory intelligence**	42	51%
Asset management	49	55%	**Risk exposure assessment**	36	43%
Quality management investigation	49	55%	**Customer profiling, targeting, and optimisation of offers**	34	41%
Smart warehousing	47	53%	**Product and service recommendation systems**	32	39%
Inventory and service parts optimisation	43	48%	**Automated customer service**	25	30%
Risk exposure assessment	39	44%	**Increase productivity and efficiency of DCs/warehouses**	21	25%
			Predictive maintenance	15	18%

Telecom/Media	# Responses	% Responses	Transport/Logistics	# Responses	% Responses
Customer profiling, targeting and optimisation of offers	67	77%	**Logistics and package delivery management**	44	67%
Product and service recommendation systems	58	67%	**New product development**	42	64%
Price optimisation	51	59%	**Inventory and service parts optimisation**	40	61%
Regulatory intelligence	50	57%	**Price optimisation**	38	58%
Automated customer service	50	57%	**Risk exposure assessment**	36	55%
New product development	48	55%	**Regulatory intelligence**	36	55%
Risk exposure assessment	42	48%	**Predictive maintenance**	36	55%
Customer scoring and/or churn mitigation	39	45%	**Supply chain optimisation**	34	52%
Network analytics and optimisation	35	40%	**Connected vehicles optimisation**	34	52%
Fraud prevention and detection	34	39%			
Scheduling optimisation	19	22%			

Utilities/Oil and Gas	# Responses	% Responses
Predictive maintenance	47	66%
Field service optimisation	45	63%
Regulatory intelligence	43	61%
Price optimisation	40	56%
Supply chain optimisation	40	56%
Energy consumption analysis and prediction	40	56%
Risk exposure assessment	39	55%
New product development	36	51%
Customer profiling, targeting and optimisation of offers	30	42%
Customer scoring and/or churn mitigation	26	37%
Inventory and service parts optimisation	9	13%
Asset management	3	4%
Quality management investigation	3	4%

Table 13 — List of Use Cases by Industry and Number of Respondents

Source: DataBench Survey, October 2018

7.5. Survey and Self-Assessment Tool on the DataBench Website

Below there is a screenshot of the link to the survey and self-assessment tool, which is available to all interested parties on the DataBench website.

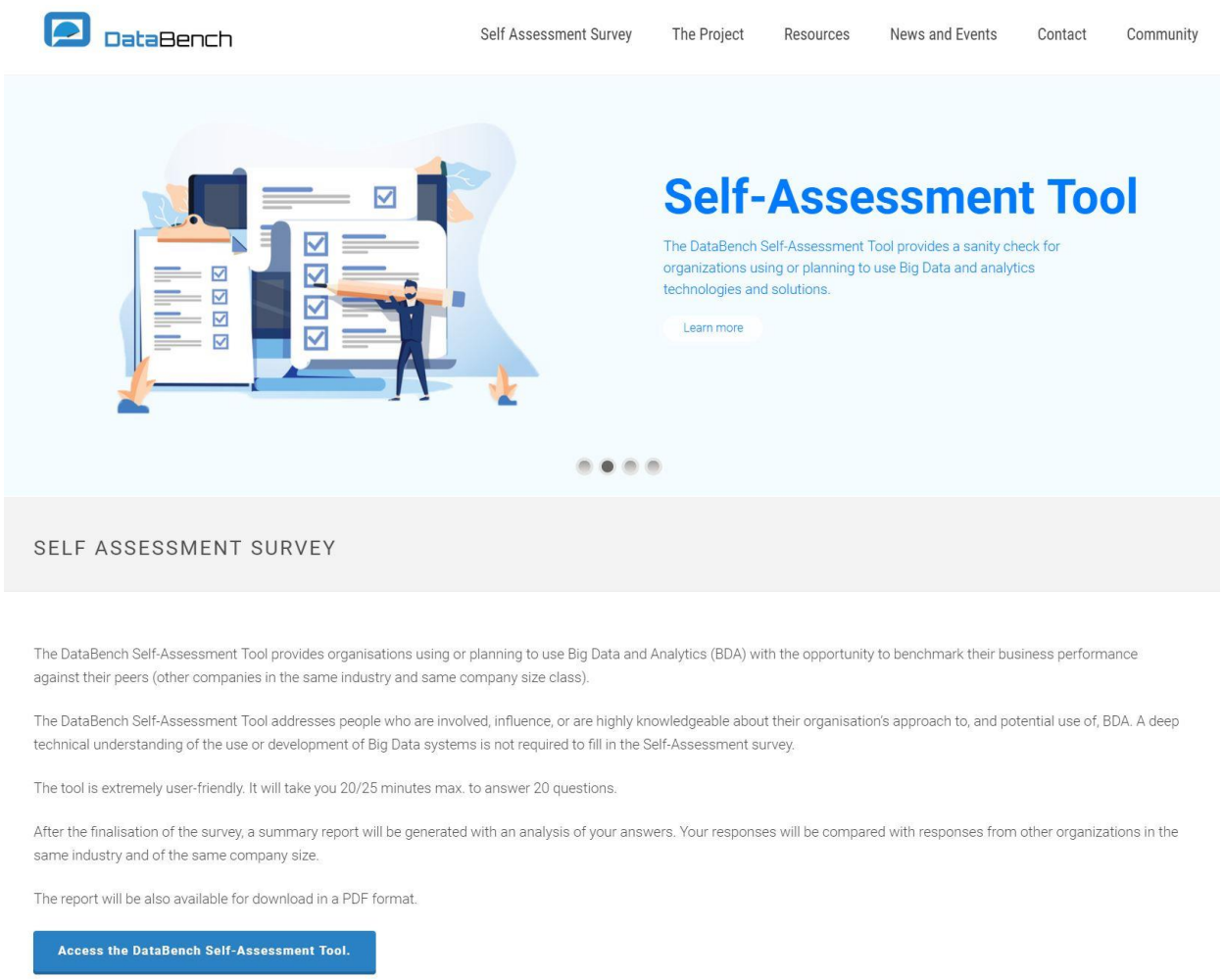


Figure 38 — Screenshot of Link to the Survey and Self-Assessment Tool on the DataBench Website

7.6. H2020 ICT Project Mapping — Classification of Trials and Pilots by Industry

The following tables provide a detailed view of the main trials and pilots implemented by H2020 ICT projects examined by DataBench. The scope of the pilots/trials, that the identification of the industry targeted by the pilots, has been examined and classified based on IDC industry classification criteria so that the projects can relate their experience to the analysis and KPIs by industry presented in this and in the previous report.

ICT	Project	Pilots and Business/Use Cases	Verticals								
			Resource Industries Agriculture	**Financial Services** Insurance Banking Securities and Investment Services	**Manufacturing** Discrete Manufacturing Process Manufacturing	**Retail & Wholesale** Retail Wholesale	**Transport & Logistics** Transportation	**Professional Services** Business/IT Services	**Telecom/Media** Telecommunication Media	**Utilities/Oil and Gas** Utilities	**Healthcare**
ICT-14	Big Data Ocean	Wave power as the next clean energy source								1	
		Maritime security and anomaly detection						1			
		Maritime protection						1			
		Vessel fault prediction					1				
	AEGIS	Automotive and road safety data									
		Smart home and assisted living									
		Insurance sector: support, warning and personal offering		1							
	euBusinessGraph	Corporate events data access service (CED)	1	1	1	1	1	1	1	1	1
		Tender discovery service (TDS)	1	1	1	1	1	1	1	1	1
		Lead generation service (Atoka +)	1	1	1	1	1	1	1	1	1
		Customer relationship service (CRM-S)	1	1	1	1	1	1	1	1	1
		Data journalism product (DJP)							1		
		Norwegian registries API service (BR-S)	1	1	1	1	1	1	1	1	1
	QROWD	Tourist network									
		Parking services									
		Traffic services					1				
	Fashion Brain	Shop the Look				1					
		Fashion trend prediction				1					
	EW-Shopp	Event-based category and brand optimisation tool			1	1					
		Weather and event-based-aware business intelligence for optimisation of campaign and resources			1	1					
		COCOS CEO workforce & campaign management optimisation		1		1	1	1	1	1	
		Measurement scout			1	1		1			
		Digital marketing campaign performance boost based on weather and event data integration	1	1	1	1	1	1	1	1	1
	DataPitch BodyPass										
	SLIPO	Geo-marketing B2C				1					
		Geo-marketing B2B	1	1	1	1	1	1	1	1	1
		Tourism									
		Transport									
	EDI										
	Lynx	Data protection	1	1	1	1	1	1	1	1	1
		Labour law	1	1	1	1	1	1	1	1	1
		Compliance assurance services, oil/gas and energy								1	
	TheyBuyForYou	Public International Trade Opportunities (PITO)	1	1	1	1	1	1	1	1	1
		Better Markets, Better Value (BMBV)	1	1	1	1	1	1	1	1	1
		Reconciliation Service (RS)	1	1	1	1	1	1	1	1	1
		Economic Data Portal (EDP)									
		Compra Pública INclusiva (COPIN)									
		Slovenian E-Publis Procurment Analysis Service (SEPPAS)									
		Vendor Intelligence Procurement Solution (VIPS)	1	1	1	1	1	1	1	1	1
	FANDANGO	Climate							1		
		Immigration							1		
		European context							1		
	Icarus	Athens					1			1	
		Basel					1			1	
		Bristol									
		Brno					1			1	
		Ljubljana					1			1	
		Madrid	1				1			1	
		Milan					1			1	
		Stuttgart					1			1	
		Thessaloniki			1		1			1	
	Cross-CPP										

Table 14 — ICT 14 Projects: Pilots and Use/Business Case Vertical Classification (IDC elaboration of H2020 data)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 780966.

ICT	Project	Pilots and Business/Use Cases	Verticals								
			Resource Industries Agriculture	**Financial Services** Insurance Banking Securities and Investment Services	**Manufacturing** Discrete Manufacturing Process Manufacturing	**Retail & Wholesale** Retail Wholesale	**Transport & Logistics** Transportation	**Professional Services** Business/IT Services	**Telecom/Media** Telecommunication Media	**Utilities/Oil and Gas** Utilities	**Healthcare**
ICT-15	DataBio	Agriculture	1								
		Fishery	1								
		Forest	1								
	TT: Transforming Transport	Highways					1				
		Rail infrastructure					1				
		Airports					1				
		Urban mobility					1				
		Vehicle connectivity					1				
		Ports					1				
		eCommerce logistics				1	1				
	BigMedilytics	Population health and chronic disease management									1
		Oncology									1
		Industrialisation of healthcare services									1
	BOOST 4.0	Volvo			1						
		Fill Your Future			1						
		Philips			1						
		Bentler			1						
		Riastone			1						
		/+GF+			1						
		Whirlpool			1						
		CRF			1						
		Gestamp			1						
		Volkswagen			1						

Table 15 — ICT 15 Projects: Pilots and Use/Business Case Vertical Classification (IDC elaboration of H2020 data)

ICT	Project	Pilots and Business/Use Cases	Verticals								
			Resource Industries Agriculture	**Financial Services** Insurance Banking Securities and Investment Services	**Manufacturing** Discrete Manufacturing Process Manufacturing	**Retail & Wholesale** Retail Wholesale	**Transport & Logistics** Transportation	**Professional Services** Business/IT Services	**Telecom/Media** Telecommunicaiton Media	**Utilities/Oil and Gas** Utilities	**Healthcare**
ICT-12	CloudButton	Genomics						1			
		Metabolomics						1			
		Geospatial					1				
	ELASTIC										
	EXA MODE										
	ExtremeEarth	Food security	1								
		Polar regions						1			
	INFORE										
	SmartDataLake										
ICT-13	MUSKETEER	Smart manufacturing			1						
		Health									1
	Safe-DEED										
ICT-16	BigDataStak	Real-time ship management					1				
		The Connected Consumer				1					
		Smart insurance		1							
	E2DATA	Health									1
		Fintech (natural language processing)		1		1			1		
		Green building infrastructure								1	
		Security and biometric recognition		1							
	Track and Know	Transport & mobility		1			1				
		Finance & insurance		1							
		Healthcare									1
	BigDataGrapes	Farm management	1								
		Natural cosmetics	1		1						
		Table and wine grapes	1								
		Wine making	1		1						
	CLASS	Intelligent traffic management					1				
		Advanced driving assistance system (ADAS)					1				
	I-BiDaaS	Accurate location prediction with high traffic and visibility							1		
		Optimisation of placement of telecommunication equipment							1		
		Employment of bots in call centres							1		
		Enhance control over third-party agencies		1							
		Advanced analysis of bank transfer payments in financial terminals		1							
		Analysis of relationships through IP address		1							
		Building of a social graph		1							
		Maintenance and monitoring of production assets			1						
		Production process of aluminium casting			1						
	Typhon										
ICT-18	SODA										
	MH-MD										
	SPECIAL	Recommendation system (PRO1)									
		Reusing telephony data for traffic alerts, optimising road layout and targeted marketing (TLABS)		1			1				
		Know-your-customer reports (TR)		1							

Table 16 — ICT 12, 13, 16 and 18 Projects: Pilots and Use/Business Case Vertical Classification (IDC elaboration of H2020 data)